

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2003, 01:06:07 ; Search time 94 Seconds
(without alignments)
5367.031 Million cell updates/sec

Title: US-09-508-997A-1
Perfect score: 1143
Sequence: 1 atggctcccttgctgcgcg.....gcaaaagtcttgccttga ll43

Scoring table: OLIGO NUC
Gapop_60.0 , Gapext 60.0

Searched: 569978 seqs, 220691566 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/prodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/prodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/prodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/prodata/2/ina/PCUTS.COMB.seq.*
6: /cgn2_6/prodata/2/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1143	100.0	1950	2	US-08-472-659-30
2	1143	100.0	1950	2	US-08-474-661-30
3	1143	100.0	1950	2	US-08-611-977-30
4	254	22.2	254	4	US-09-016-434-658
5	230	20.1	316	4	US-09-016-434-824
6	121	10.6	600	4	US-09-702-705-233
7	121	10.6	600	4	US-09-736-457-233
8	25	2.2	37	2	US-08-472-659-32
9	25	2.2	37	2	US-08-474-661-32
10	25	2.2	37	2	US-08-611-977-32
11	23	2.0	23	2	US-08-472-659-14
12	23	2.0	23	2	US-08-474-661-14
13	23	2.0	23	2	US-08-611-977-14
14	21	1.8	21	2	US-08-472-659-15
15	21	1.8	21	2	US-08-472-659-16
16	21	1.8	21	2	US-08-472-659-27
17	21	1.8	21	2	US-08-472-659-28
18	21	1.8	21	2	US-08-472-659-29
19	21	1.8	21	2	US-08-474-661-15
20	21	1.8	21	2	US-08-474-661-16
21	21	1.8	21	2	US-08-474-661-27
22	21	1.8	21	2	US-08-474-661-28
23	21	1.8	21	2	US-08-474-661-29
24	21	1.8	21	2	US-08-611-977-15
25	21	1.8	21	2	US-08-611-977-16
26	21	1.8	21	2	US-08-611-977-27
27	21	1.8	21	2	US-08-611-977-28

C	28	21	1.8	21	2	US-08-611-977-29	Sequence 29, Appl
	29	20	1.7	83450	4	US-09-811-469-3	Sequence 3, Appl
C	30	19	1.7	3001	4	US-09-539-333D-150	Sequence 150, App
	31	18	1.6	699	3	US-09-328-111-291	Sequence 291, App
C	32	18	1.6	798	4	US-09-328-352-2221	Sequence 2221, Ap
	33	18	1.6	2064	4	US-09-482-273-57	Sequence 57, Appl
C	34	18	1.6	169984	4	US-09-676-610B-24	Sequence 24, Appl
	35	18	1.6	137496	4	US-09-877-177A-10	Sequence 10, Appl
C	36	18	1.6	1664976	4	US-08-516-421B-1	Sequence 1, Appl
	37	17	1.5	27	2	US-07-594-921C-3	Sequence 3, Appl
	38	17	1.5	27	4	US-09-197-948-3	Sequence 11, Appl
C	39	17	1.5	124	2	US-07-594-921C-11	Sequence 11, Appl
	40	17	1.5	124	4	US-09-197-948-11	Sequence 11, Appl
C	41	17	1.5	237	4	US-09-016-434-460	Sequence 460, App
	42	17	1.5	248	2	US-07-594-921C-10	Sequence 10, Appl
C	43	17	1.5	248	4	US-09-197-948-10	Sequence 10, Appl
	44	17	1.5	521	4	US-08-858-207A-252	Sequence 252, App
	45	17	1.5	658	4	US-09-641-638-402	Sequence 402, App

ALIGNMENTS

RESULT 1
US-08-472-659-30
; Sequence 30, Application US/08472659
; Patent No. 5831030
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Fuyuki
; APPLICANT: TSUJIOKA, No. 5831030uo
; APPLICANT: NAKAZATO, Hiroshi
; APPLICANT: MURA, Kenju
; APPLICANT: ISHIDA, No. 5831030uhiro
; APPLICANT: KURIHARA, Tatsuya
; APPLICANT: YAMAUCHI, Kozo
; APPLICANT: YAMAGUCHI, No. 5831030oml
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,659
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5-067339
; FILING DATE: 04-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-248
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:

Query Match	100.08;	Score 1143;	DB 2;	Length 1950;		
Best Local Similarity	100.04;	Pred. No. 0;				
Matches 1143;	Conservative	0;	Mismatches	0; Indels	0; Gaps	0;
QY	1	ATGGCTCCCTCTGCTGCAGCAAAATGAGAGATTTTTCCTTCAACCTCTTCAGAGAGATGGAT	50			
Db	74	ATGGCTCCCTCTGCTGCAGCAAAATGAGAGATTTTTCCTTCAACCTCTTCAGAGAGATGGAT	133			
QY	61	GACAAATCAGGAANAATGGAAATATGCTTTCTTCCTCTCGAGCCTCTTGCTGCCCTGCCC	120			
Db	134	GACAAATCAGGAANAATGGAAATATGCTTTCTTCCTCTCGAGCCTCTTGCTGCCCTGCCC	193			
QY	121	CTGGTCCGCTTGGCGCTCMAGATGACTCCCTCTCTCAGATGATTAAGTTGCTTCAATGTT	180			
Db	194	CTGGTCCGCTTGGCGCTCMAGATGACTCCCTCTCTCAGATGATTAAGTTGCTTCAATGTT	253			
QY	181	AACACTGGCTCAGGATATCGGAACCTTTCTAATGATGTCGGGCTCCAGTCTCAGACTG	240			
Db	254	AACACTGGCTCAGGATATCGGAACCTTTCTAATGATGTCGGGCTCCAGTCTCAGACTG	313			
QY	241	AAAAGAGTTTTTCTGATATAAATGCAATCCCAAGAGATTATGATCTCAGCATTTGTAAT	300			
Db	314	AAAAGAGTTTTTCTGATATAAATGCAATCCCAAGAGATTATGATCTCAGCATTTGTAAT	373			
QY	301	GGGCTTTTGTGTAAGGATGATGGCTTTTCATPAAGACTACATTGAGTGTGCCGAAAA	360			
Db	374	GGGCTTTTGTGTAAGGATGATGGCTTTTCATPAAGACTACATTGAGTGTGCCGAAAA	433			
QY	361	TTAPAGATGCCAAAGTGGAGCGAGTTGCACTTTTACGATCAATTTAGAGACACTAGACGT	420			
Db	434	TTAPAGATGCCAAAGTGGAGCGAGTTGCACTTTTACGATCAATTTAGAGACACTAGACGT	493			
QY	421	AAATATTAATGAGGGTTGAAATGAAACACATGCGCAAAATCAAGAACCTGATTTGGTGAA	480			
Db	494	AAATATTAATGAGGGTTGAAATGAAACACATGCGCAAAATCAAGAACCTGATTTGGTGAA	553			
QY	481	GGTGCCATAAGCTCATCTGCTGTAAATGCTGGTGGAATGCTGTGTACTTCAAGGCAAG	540			
Db	554	GGTGCCATAAGCTCATCTGCTGTAAATGCTGGTGGAATGCTGTGTACTTCAAGGCAAG	613			
QY	541	TGGCAATTCAGCCTTCAACAGAGCGAAACCAATAATGCGCATTTCAATCTCCCAAGTGC	600			
Db	614	TGGCAATTCAGCCTTCAACAGAGCGAAACCAATAATGCGCATTTCAATCTCCCAAGTGC	673			
QY	601	TCGTGGAGGACGTGGCCATGATGCATCAGGAACGGAGTTCAATTTGCTGTATTGAG	650			
Db	674	TCGTGGAGGACGTGGCCATGATGCATCAGGAACGGAGTTCAATTTGCTGTATTGAG	733			
QY	661	GACCCATCAATGAAGATCTTTGAGCTTCAGATCAATGGTGGCAATPAACATGTACCTTCG	720			
Db	734	GACCCATCAATGAAGATCTTTGAGCTTCAGATCAATGGTGGCAATPAACATGTACCTTCG	793			
QY	721	CTGGCTGGAATGACCTCTTGAAATTTGAACCAAACTGACCTTTTCAATCTTAATGGAA	780			
Db	794	CTGGCTGGAATGACCTCTTGAAATTTGAACCAAACTGACCTTTTCAATCTTAATGGAA	853			

QY	781	TGGACCAATCCAAGGCGAATGACCTCTTAGTATGTTGAGGTATTTTTCCTCAGTTCAAG	840
Db	854	TGGACCAATCCAAGGCGAATGACCTCTTAGTATGTTGAGGTATTTTTCCTCAGTTCAAG	913
QY	841	ATAGAGAGAAATATGAAATGAACAATATTTTGAGAGCCCTAGGCTGAAAGATATCTTT	900
Db	914	ATAGAGAGAAATATGAAATGAACAATATTTTGAGAGCCCTAGGCTGAAAGATATCTTT	973
QY	901	GATGAATCCAAGCAGATCTCTTGGGATTCCTTGGGGGCTCGTCTGTATATCAAGG	960
Db	974	GATGAATCCAAGCAGATCTCTTGGGATTCCTTGGGGGCTCGTCTGTATATCAAGG	1033
QY	961	ATGATGCGCAAACTTACATAGAGTCACTGAGAGGCGCACCGAGGCTACTCTGTCACA	1020
Db	1034	ATGATGCGCAAACTTACATAGAGTCACTGAGAGGCGCACCGAGGCTACTCTGTCACA	1093
QY	1021	GGAGATTAATATTGTAGAAAAGCAATCCCTCAGTCCAGCTGTTTAGAGCTGACCAACCA	1080
Db	1094	GGAGATTAATATTGTAGAAAAGCAATCCCTCAGTCCAGCTGTTTAGAGCTGACCAACCA	1153
QY	1081	TTCTATTCTGTTATGAGGAAGTGAATCATCATCTTATTCAGTGCAGAAAGTTCTTGCCCT	1140
Db	1154	TTCTATTCTGTTATGAGGAAGTGAATCATCATCTTATTCAGTGCAGAAAGTTCTTGCCCT	1213
QY	1141	TGA 1143	
Db	1214	TGA 1216	
		RESULT 2	
		US-08-474-661-30	
		Sequence 30, Application US/08474661	
		Patent No. 5874253	
		GENERAL INFORMATION:	
		APPLICANT: TSUJIMOTO, Masafumi	
		APPLICANT: IWASA, Fuyuki	
		APPLICANT: TSURUOKA, No. 5874253uo	
		APPLICANT: NAKAZATO, Hiroshi	
		APPLICANT: MURA, Kenju	
		APPLICANT: ISHIDA, No. 5874253uhiro	
		APPLICANT: KURIHARA, Tatsuya	
		APPLICANT: YAMAICHI, Kozo	
		APPLICANT: YAMAGUCHI, No. 5874253omi	
		TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR	
		NUMBER OF SEQUENCES: 34	
		CORRESPONDENCE ADDRESS:	
		ADDRESSER: Burns, Doane, Swecker & Mathis	
		STREET: George Mason Bldg., Washington & Prince Sts.	
		CITY: Alexandria	
		STATE: Virginia	
		COUNTRY: United States	
		ZIP: 22313-1404	
		COMPUTER READABLE FORM:	
		MEDIUM TYPE: Floppy disk	
		COMPUTER: IBM PC compatible	
		OPERATING SYSTEM: PC-DOS/MS-DOS	
		SOFTWARE: Patentin Release #1.0, Version #1.25	
		CURRENT APPLICATION DATA:	
		APPLICATION NUMBER: US/08/474,661	
		FILING DATE: 07-JUN-1995	
		CLASSIFICATION: 435	
		PRIOR APPLICATION DATA:	
		APPLICATION NUMBER: US 08/091,028	
		FILING DATE: 14-JUL-1993	
		APPLICATION NUMBER: JP 4-212305	
		FILING DATE: 17-JUL-1992	
		PRIOR APPLICATION DATA:	
		APPLICATION NUMBER: JP 5-067339	
		FILING DATE: 04-MAR-1993	
		ATTORNEY/AGENT INFORMATION:	
		NAME: SEA, TERESA STANEK	
		REGISTRATION NUMBER: 30,427	

REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 1950 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Peptide
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
STRAIN: A431
FEATURE:
NAME/KEY: misc feature
LOCATION: 1
OTHER INFORMATION: /note= "DNA coding for human megakaryocyte differentiation factor."
FEATURE:
NAME/KEY: CDS
LOCATION: 74..1217
US-08-474-661-30

Query Match 100.0%; Score 1143; DB 2; Length 1950;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTCCCTGCTGTCAGCAATGCAAGTGGTTCCTCAACCTGTCAGAGAGATGGAT 60
DB 74 ATGGCTCCCTGCTGTCAGCAATGCAAGTGGTTCCTCAACCTGTCAGAGAGATGGAT 133
QY 61 GACATCAAGAGAAATGGAATGTGTTCTTTCTCTCTGAGGCTCTTGCTGCGCTGGCC 120
DB 134 GACATCAAGAGAAATGGAATGTGTTCTTTCTCTCTGAGGCTCTTGCTGCGCTGGCC 193
QY 121 CTGGTCCGCTGGGCGCTCAAGTACCTCTCTCAATAGTCAGTCCAGTTCAGATGATGCTCATGTT 180
DB 194 CTGGTCCGCTGGGCGCTCAAGTACCTCTCTCAATAGTCAGTTCAGATGATGCTCATGTT 253
QY 181 AACCTGCTCAGGATGCAAACTCTCTAATAGTCAGTCCAGGCTCCAGTCTCAACTG 240
DB 254 AACCTGCTCAGGATGCAAACTCTCTAATAGTCAGTCCAGGCTCCAGTCTCAACTG 313
QY 241 AAAAGAGTTTTTCTGATATAAATGCAATCCCAAGGATATGATCTCAGCATGTGAT 300
DB 314 AAAAGAGTTTTTCTGATATAAATGCAATCCCAAGGATATGATCTCAGCATGTGAT 373
QY 301 GGGCTTTTCTGAAAAAGTGTATGCTTTTCAATAGGACTACATGATGCTGCGGAAAA 360
DB 374 GGGCTTTTCTGAAAAAGTGTATGCTTTTCAATAGGACTACATGATGCTGCGGAAAA 433
QY 361 TTATAGATGCCAAAGTGGAGGAGTGAATGCTTTTCAATAGGACTACATGATGCTGCGGAAAA 420
DB 434 TTATAGATGCCAAAGTGGAGGAGTGAATGCTTTTCAATAGGACTACATGATGCTGCGGAAAA 493
QY 421 AATATTAATAGTGGTTGAAATGAAACATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 480
DB 494 AATATTAATAGTGGTTGAAATGAAACATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 553
QY 481 GGTGGCATAGCTCATCTGCTGTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
DB 554 GGTGGCATAGCTCATCTGCTGTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 613
QY 541 TGGCAATCAAGCTTCAACCAAGAGCAATCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 600
DB 614 TGGCAATCAAGCTTCAACCAAGAGCAATCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 673
QY 601 TCTGGGAAGCAGTCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 674 TCTGGGAAGCAGTCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 733
QY 661 GAACCATCAATGAGATTTCTTGAGCTCAGATACATGCTGGGCAATACATGATGATGCTGCTG 720

DB 734 GAACCATCAATGAGATTTCTTGAGCTCAGATACATGCTGGGCAATACATGATGCTGCTG 793
QY 721 CTGCTTGAGATGACCTCTCTGAAATGAAACAACTGACCTTTCAGATCTAATGGA 780
DB 794 CTGCTTGAGATGACCTCTCTGAAATGAAACAACTGACCTTTCAGATCTAATGGA 853
QY 781 TGGACCAATCCAGGCAATGACCTCTAAGTATGCTGAGGTATTTTTCCTCAGTTCAAG 840
DB 854 TGGACCAATCCAGGCAATGACCTCTAAGTATGCTGAGGTATTTTTCCTCAGTTCAAG 913
QY 841 ATGAGAGCAATTTGAAATGAAACAACTGACCTTTCAGATCTAATGGA 900
DB 914 ATGAGAGCAATTTGAAATGAAACAACTGACCTTTCAGATCTAATGGA 973
QY 901 GATGAATCCAAAGCAGATCTCTCTGAGGTGCTTGGGGGGGCTGCTCTATATATCAAG 960
DB 974 GATGAATCCAAAGCAGATCTCTCTGAGGTGCTTGGGGGGGCTGCTCTATATATCAAG 1033
QY 961 ATGATGCAATCTTACATGAGGTCTACTGAGGAGGCAACGAGGCTACTGCTGCA 1020
DB 1034 ATGATGCAATCTTACATGAGGTCTACTGAGGAGGCAACGAGGCTACTGCTGCA 1093
QY 1021 GGAAGTAAATTTAGAAAAAGCACTCCCTCAGTCCACGCTGTTTAGAGCTGACACCCA 1080
DB 1094 GGAAGTAAATTTAGAAAAAGCACTCCCTCAGTCCACGCTGTTTAGAGCTGACACCCA 1153
QY 1081 TTCTATTTGTTATCAGAGGATGATCATCTTATTCAGTGGCAAGATTTCTGCTT 1140
DB 1154 TTCTATTTGTTATCAGAGGATGATCATCTTATTCAGTGGCAAGATTTCTGCTT 1213
QY 1141 TGA 1143
DB 1214 TGA 1216

RESULT 3
US-08-611-977-30
Sequence 30, Application US/08611977
Patent No. 5972886
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSUBOIUCHI, No. 5972886uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5972886uhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAUCHI, Kozo
APPLICANT: YAMAUCHI, No. 5972886omi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/611,977
FILING DATE: 06-MAR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305

FILED DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 1950 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
STRAIN: A431
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1
OTHER INFORMATION: /note= "DNA coding for human megakaryocyte differentiation factor."
FEATURE:
NAME/KEY: CDS
LOCATION: 74..1217
US-08-611-977-30

Query Match 100.0%; Score 1143; DB 2; Length 1950;
Best Local Similarity 100.0%; Pred No. 0;
Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCGCTCCCTTCTCCAGCAAAATGACAGATTTTGGCTTCAACCTGTTTCAGAGATGGAT 60
DB 74 ATGCGCTCCCTTCTCCAGCAAAATGACAGATTTTGGCTTCAACCTGTTTCAGAGATGGAT 133
QY 61 GACATCAGAGAAATGGAATGTCTTTCTCTCTGAGCGCTTCGCTGCCCTGCC 120
DB 134 GACAATCAAGAAATGGAATGTCTTTCTCTCTGAGCGCTTCGCTGCCCTGCC 193
QY 121 CTGTCCTCCCTTGGCGCTCAGATGATCCCTCTCTCAGATGATGATGATGATGAT 180
DB 194 CTGTCCTCCCTTGGCGCTCAGATGATCCCTCTCTCAGATGATGATGATGATGAT 253
QY 181 AACACTGCTCAGATATGGAATCTTTCTTAATAGTCAGTCCAGGCTCCAGTCAACTG 240
DB 254 AACACTGCTCAGATATGGAATCTTTCTTAATAGTCAGTCCAGGCTCCAGTCAACTG 313
QY 241 AAAAGATTTTCTGATATAATGATGATGATGATGATGATGATGATGATGATGATGAT 300
DB 314 AAAAGATTTTCTGATATAATGATGATGATGATGATGATGATGATGATGATGATGAT 373
QY 301 GGCCTTTTGTGAAAAGTGTATGCTTTTCAATGAGTCAATGATGATGATGATGATGAT 360
DB 374 GGCCTTTTGTGAAAAGTGTATGCTTTTCAATGAGTCAATGATGATGATGATGATGAT 433
QY 361 TTATAGATGCAAGTGGAGCGAGTGTGACTTTACGATCATTTAGAGACATAGAGT 420
DB 434 TTATAGATGCAAGTGGAGCGAGTGTGACTTTACGATCATTTAGAGACATAGAGT 493
QY 421 AATATTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
DB 494 AATATTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 553
QY 481 GGTGGCATAGCTCATCTGCTGTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
DB 554 GGTGGCATAGCTCATCTGCTGTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 613
QY 541 TGGCAATCAGCTTCCACAGAGCGGAACCAATTAATGCAATTCATATCCCAAGTCC 600

DB 614 TGGCAATCAGCTTCCACAGAGCGGAACCAATTAATGCAATTCATATCCCAAGTCC 573
QY 601 TCTGGAGGCGAGTGGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 560
DB 674 TCTGGAGGCGAGTGGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 733
QY 661 GACCCATCAATGAAGTCTTGGAGTCTCAGATACATGATGATGATGATGATGATGATGATGAT 720
DB 734 GACCCATCAATGAAGTCTTGGAGTCTCAGATACATGATGATGATGATGATGATGATGATGAT 793
QY 721 CTGCTCAGAGTACCTCTCTGAAATGAAACAACTGACCTTTTCTCAGATCTAATGAA 780
DB 794 CTGCTCAGAGTACCTCTCTGAAATGAAACAACTGACCTTTTCTCAGATCTAATGAA 853
QY 781 TGGACCAATCCAAAGCGAATGACCTCTAAGTATGTTGAGTATTTTCTCAGTCTCAAG 840
DB 854 TGGACCAATCCAAAGCGAATGACCTCTAAGTATGTTGAGTATTTTCTCAGTCTCAAG 913
QY 841 ATAGAGAGAAATTTATGAATGAAACAACTGACCTTTTCTCAGATCTAATGAA 900
DB 914 ATAGAGAGAAATTTATGAATGAAACAACTGACCTTTTCTCAGATCTAATGAA 973
QY 901 GATGAATCCAAAGCGAATGACCTCTTGGGATGCTTGGGGGCTCGTCTGATATATCAAG 960
DB 974 GATGAATCCAAAGCGAATGACCTCTTGGGATGCTTGGGGGCTCGTCTGATATATCAAG 1033
QY 961 ATGATGCAAAATCTTAAATGAGTCACTGAGGAGGACCGAGGCTACTGCTGCCACA 1020
DB 1034 ATGATGCAAAATCTTAAATGAGTCACTGAGGAGGACCGAGGCTACTGCTGCCACA 1093
QY 1021 GGAAGTATATTTAGAAAGCACTCCCTCAGTCCACGCTTTTAGAGTGACCAACCCA 1080
DB 1094 GGAAGTATATTTAGAAAGCACTCCCTCAGTCCACGCTTTTAGAGTGACCAACCCA 1153
QY 1081 TTCTCTATTTGTTATCAGAGGATGATCATCTTATTCAGTGGCAAGTTCTTGGCTT 1140
DB 1154 TTCTCTATTTGTTATCAGAGGATGATCATCTTATTCAGTGGCAAGTTCTTGGCTT 1213
QY 1141 TGA 1143
DB 1214 TGA 1216

RESULT 4
US-09-016-434-658
Sequence 658, Application US/09016434
Patent No. 6500938
GENERAL INFORMATION:
APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:

Query Match 20.1%; Score 230; DB 4; Length 316;
Best Local Similarity 99.6%; Pred. No. 6.3e-110;
Matches 280; Conservative 0; Mismatches 1; Indels

Qy	822	ATTTTTCCTCAGTTCACAGATGAGAGAGAAATTAAGAAATGAAACAATATTTGACAGCCCT	888
Db	1	ATTTTTCCTCAGTTCACAGATGAGAGAGAAATTAAGAAATGAAACAATATTTGACAGCCCT	60
Qy	882	AGGGCTGAAAGATATCTTTTGATGAATCCAAACAGATCTCTCTGGGATGGTTTCGGGGGG	941
Db	61	AGGGCTGAAAGATATCTTTTGATGAATCCAAACAGATCTCTCTGGGATGGTTTCGGGGGG	120
Qy	942	TGCTCTGTATATATCAAGGATGATGCACAAATCTTACATAGAGGTCACTGAGGAGGGCAC	1001
Db	121	TGCTCTGTATATATCAAGGATGATGCACAAATCTTACATAGAGGTCACTGAGGAGGGCAC	180
Qy	1002	CGAGGCTACTGCTGGCCACAGGAAGTAATTTCTGAAAGCAACTCCCTCAGTCCAGCT	1061
Db	181	CGAGGCTACTGCTGGCCACAGGAAGTAATTTCTGAAAGCAACTCCCTCAGTCCAGCT	240
Qy	1062	GTTTAGCTGACCAACCCATTCCTATTTGTTATCAGGAAG	1102
Db	241	GTTTAGCTGACCAACCCATTCCTATTTGTTATCAGGAAG	281

```

RESULT 6
US-09-702-705-233/c
; Sequence 233, Application US/09702705
; Patent No. 6504010
;
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Barrick
; APPLICANT: Retter, Marc
; APPLICANT: Mammion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 233
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-233

Query Match          10.6%; Score 121; DB 4; Length 600;
Best Local Similarity 100.0%; Pred. No. 3.7e-53;
Matches 12; Conservative 0; Mismatches 0; Indels 0;

```

```
Query Match: 10.6%; Score 121; DB 4; Length 600;
Best Local Similarity 100.0%; Pred. No. 3.7e-53;
Matches 121; Conservative 0; Mismatches 0; Indels
```

Mon Dec 15 08:51:17 2003

us-09-508-997a-1.rni

TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR

NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,659
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536

PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA: JP 5-067339
FILING DATE: 04-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993

ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-248
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-472-659-32

Query Match 2.2%; Score 25; DB 2; Length 37;
Best Local Similarity 100.0%; Pred. No. 0.0032;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCTCCCTTCCTGCGCAATG 25
Db 13 ATGGCTCCCTTCCTGCGCAATG 37

RESULT 9

US-08-474-661-32
Sequence 32, Application US/08474661
Patent No. 5874253
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURUOKA, No. 5874253uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5874253uhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAUCHI, Kozo
APPLICANT: YAMAGUCHI, No. 5874253omi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSES: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia

QY 1023 AAGTATATTGTAGAAAGCACTCCCTCAGTCAGCGCTTTAGAGCTGACCAACCAATT 1082
Db 600 AAGTATATTGTAGAAAGCACTCCCTCAGTCAGCGCTTTAGAGCTGACCAACCAATT 541
QY 1083 CCTATTGTGTATCAGGAAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGCCCTTG 1142
Db 540 CCTATTGTGTATCAGGAAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGCCCTTG 481

RESULT 7

US-09-736-457-233/c
Sequence 233, Application US/09736457
Patent No. 6509448
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
APPLICANT: Wang, Aijun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C15
CURRENT APPLICATION NUMBER: US/09/736,457
CURRENT FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 1864
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 233
LENGTH: 600
TYPE: DNA
ORGANISM: Homo sapien
US-09-736-457-233

Query Match 10.6%; Score 121; DB 4; Length 600;
Best Local Similarity 100.0%; Pred. No. 3.7e-53;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1023 AAGTATATTGTAGAAAGCACTCCCTCAGTCAGCGCTTTAGAGCTGACCAACCAATT 1082
Db 600 AAGTATATTGTAGAAAGCACTCCCTCAGTCAGCGCTTTAGAGCTGACCAACCAATT 541
QY 1083 CCTATTGTGTATCAGGAAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGCCCTTG 1142
Db 540 CCTATTGTGTATCAGGAAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGCCCTTG 481
QY 1143 A 1143
Db 480 A 480

RESULT 8

US-08-472-659-32
Sequence 32, Application US/08472659
Patent No. 5831030
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURUOKA, No. 5831030uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5831030uhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAUCHI, Kozo
APPLICANT: YAMAGUCHI, No. 5831030omi

COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,661
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: REA, TERESA STANEK
REGISTRATION NUMBER: 30,427
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-474-661-32

Query Match 2.2%; Score 25; DB 2; Length 37;
Best Local Similarity 100.0%; Pred. No. 0.0032;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTCCCTTGCTGCAGCAATG 25
Db 13 ATGGCTCCCTTGCTGCAGCAATG 37

RESULT 10
US-08-611-977-32
Sequence 32, Application US/08611977
Patent No. 5972886
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSUROOKA, No. 5972886u
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5972886uhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAUCHI, Kozo
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,661
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-611-977-32

Query Match 2.2%; Score 25; DB 2; Length 37;
Best Local Similarity 100.0%; Pred. No. 0.0032;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTCCCTTGCTGCAGCAATG 25
Db 13 ATGGCTCCCTTGCTGCAGCAATG 37

RESULT 11
US-08-472-659-14
Sequence 14, Application US/08472659
Patent No. 5831030
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSUROOKA, No. 5831030u
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5831030uhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAUCHI, Kozo
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,659
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:

us-09-508-997a-1.rni

Mon Dec 15 08:51:17 2003

```

; APPLICATION NUMBER: JP 5-067339
; FILING DATE: 04-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-248
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-472-659-14

```

```

Query Match 2.0%; Score 23; DB 2; Length 23;
Best Local Similarity 100.0%; Pred.No. 0.034;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 791 CAAGCGGAATGACCTCTAAGTAT 813
DB 1 CAAGCGGAATGACCTCTAAGTAT 23

```

```

RESULT 12
US-08-474-661-14
; Sequence 14, Application US/08474661
; Patent No. 5874253
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Fuyuki
; APPLICANT: TSUJOKA, No. 5874253uo
; APPLICANT: NAKAZATO, Hiroschi
; APPLICANT: MIURA, Kenju
; APPLICANT: ISHIDA, No. 5874253uhiro
; APPLICANT: KURIHARA, Tatsuya
; APPLICANT: YAMAICHI, Kozo
; APPLICANT: YAMAGUCHI, No. 5874253omi
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: George Mason Bldg., Washington & Prince Sts.
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,661
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5-067339
; FILING DATE: 04-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: REA, TERESA STANEX
; REGISTRATION NUMBER: 30,427

```

```

; REFERENCE/DOCKET NUMBER: 001560-204
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-6620
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-474-661-14

```

```

Query Match 2.0%; Score 23; DB 2; Length 23;
Best Local Similarity 100.0%; Pred.No. 0.034;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 791 CAAGCGGAATGACCTCTAAGTAT 813
DB 1 CAAGCGGAATGACCTCTAAGTAT 23

```

```

RESULT 13
US-08-611-977-14
; Sequence 14, Application US/08611977
; Patent No. 5972886
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Fuyuki
; APPLICANT: TSUJOKA, No. 5972886uo
; APPLICANT: NAKAZATO, Hiroschi
; APPLICANT: MIURA, Kenju
; APPLICANT: ISHIDA, No. 5972886uhiro
; APPLICANT: KURIHARA, Tatsuya
; APPLICANT: YAMAICHI, Kozo
; APPLICANT: YAMAGUCHI, No. 5972886omi
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611,977
; FILING DATE: 06-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-067339
; FILING DATE: 04-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-204
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 base pairs

```

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-611-977-14

Query Match 2.0%; Score 23; DB 2; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.034;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 791 CAAGCGGAATGACCTCTAAGTAT 813
DB 1 CAAGCGGAATGACCTCTAAGTAT 23

RESULT 14
US-08-472-659-15/c
Sequence 15, Application US/08472659
Patent No. 5831030
GENERAL INFORMATION:
APPLICANT: TSUTIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURUOKA, No. 5831030uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: IGHIDA, No. 5831030uhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. 5831030omi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,659
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-248
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-472-659-15

Query Match 1.8%; Score 21; DB 2; Length 21;

Best Local Similarity 100.0%; Pred. No. 0.37;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 919 CTCCTCTGGGATTCCTTCGGGG 939
DB 21 CTCCTCTGGGATTCCTTCGGGG 1

RESULT 15
US-08-472-659-16/c
Sequence 16, Application US/08472659
Patent No. 5831030
GENERAL INFORMATION:
APPLICANT: TSUTIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURUOKA, No. 5831030uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: IGHIDA, No. 5831030uhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. 5831030omi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,659
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-248
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-472-659-16

Query Match 1.8%; Score 21; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.37;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 709 ATGTACCTTCTGCTGCTGAG 729
DB 21 ATGTACCTTCTGCTGCTGAG 1

us-09-508-997a-1.rni

Mon Dec 15 08:51:17 2003

Search completed: December 14, 2003, 03:21:09
Job time : 101 secs

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

CM nucleic - nucleic search, using sw model

Run on: December 14, 2003, 03:21:22 ; Search time 4369 Seconds
(without alignments)
10702.610 Million cell updates/sec

Title: US-09-508-997a-1

Perfect score: 1143

Sequence: 1 atggctcccttgcgtgcagc.....gcaagtttttgccttga 1143

Scoring table: OLIQO NUC
Gapop 60.0 , Gapext 60.0

Searched: 2888711 seqs, 2045481386 residues

Word size: 0

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.on.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pin.*

35: em.htg.rod.*

36: em.htg.mam.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

Result No.	Score	Query Match	Length	DB	ID	Description	
1	1143	100.0	1143	6	BD094240	BD094240 Model ani	
2	1143	100.0	1950	6	AR052445	AR052445 Sequence	
3	1143	100.0	1950	6	AR082433	AR082433 Sequence	
4	1143	100.0	1950	6	E08396	E08396 cDNA encodi	
5	1143	100.0	1950	9	D88575	D88575 Homo sapien	
6	1143	100.0	2249	9	AF027866	AF027866 Homo sapi	
C	7	399	34.9	157284	9	AC072051	AC072051 Homo sapi
8	399	34.9	189092	2	AC072524	AC072524 Sequence	
9	254	22.2	254	6	AR270095	AR270095 Sequence	
10	230	20.1	316	6	AR270261	AR270261 Sequence	
C	11	121	10.6	600	6	AR272521	AR272521 Sequence
C	12	121	10.6	600	6	AR276102	AR276102 Sequence
C	13	121	10.6	600	6	AX062606	AX062606 Sequence
C	14	121	10.6	600	6	AX367523	AX367523 Sequence
15	31	2.7	1143	10	AF105329	AF105329 Rattus no	
16	31	2.7	1147	6	BD094242	BD094242 Model ani	
17	31	2.7	1229	6	BD094241	BD094241 Model ani	
18	31	2.7	1229	6	BD096261	BD096261 Method fo	
19	31	2.7	1386	10	AF105328	AF105328 Mus muscu	
20	31	2.7	121141	10	AC125314	AC125314 Mus muscu	
21	30	2.6	30	6	BD094249	BD094249 Model ani	
22	30	2.6	44	6	BD094243	BD094243 Model ani	
23	29	2.5	66599	2	AC129921	AC129921 Mus muscu	
24	29	2.5	199669	10	AC103453	AC103453 Rattus no	
C	25	29	2.5	219071	2	AC133259	AC133259 Rattus no
C	26	2.3	26	6	BD094244	BD094244 Model ani	
27	25	2.2	37	6	AR052447	AR052447 Sequence	
28	25	2.2	37	6	AR082435	AR082435 Sequence	
29	25	2.2	37	6	E08397	E08397 PCR primer	
30	23	2.0	23	6	AR052429	AR052429 Sequence	
31	23	2.0	23	6	AR082417	AR082417 Sequence	
32	23	2.0	142127	10	AC122521	AC122521 Mus muscu	
C	33	23	2.0	179401	2	BX294444	BX294444 Danio rer
34	23	2.0	198515	10	AF129005	AF129005 Mus muscu	
C	35	23	2.0	285727	2	AC025117	AC025117 Mus muscu
36	22	1.9	19950	3	AF267745	AF267746 Drosophil	
C	37	22	1.9	51602	2	AC129261	AC129261 Rattus no
C	38	22	1.9	89248	2	AC021919	AC021919 Homo sapi
C	39	22	1.9	144430	9	AC091155	AC091155 Homo sapi
40	22	1.9	168509	9	AC009837	AC009837 Homo sapi	
C	41	22	1.9	171627	2	AC126924	AC126924 Sus scrof
C	42	22	1.9	172195	9	AC093862	AC093862 Homo sapi
C	43	21	1.8	21	6	AR052430	AR052430 Sequence
C	44	21	1.8	21	6	AR052431	AR052431 Sequence
C	45	21	1.8	21	6	AR052442	AR052442 Sequence

ALIGNMENTS

RESULT 1	BD094240	BD094240	1143 bp	DNA	linear	PAT 27-AUG-2002
LOCUS	BD094240	Model animals of nephropathy proliferating mesangial cells.				
DEFINITION	BD094240	Model animals of nephropathy proliferating mesangial cells.				
ACCESSION	BD094240	Model animals of nephropathy proliferating mesangial cells.				
VERSION	BD094240.1	GI:22639828				
KEYWORDS	WO 0124628-A/1.					
SOURCE	WO 0124628-A/1.					
ORGANISM	Homo sapiens (human)					
REFERENCE	1 (bases 1 to 1143)					
AUTHORS	Miyata,T.					
TITLE	Model animals of nephropathy proliferating mesangial cells					
JOURNAL	Patent: WO 0124628-A 1 12-APR-2001;					

COMMENT
TOSHIO MIYATA, KIYOSHI KUROKAWA
CS Homo sapiens (human)
FN WO 0124628-A/1
PD 12-APR-2001
PF 06-OCT-2000 WO 2000JP006988
PR 06-OCT-1999 JP 99P 285736
PI TOSHIO MIYATA
PC A01K67/027, A61P13/12, A61K45/00, C12N15/12, C12N15/85, G01N33/15,
CC G01N33/50

Key Location/Qualifiers
CDs Location/Qualifiers
i. .1143
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
source

BASE COUNT 336 a 230 c 261 g 316 t

Query Match 100.0%; Score 1143; DB 6; Length 1143;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCCCTCCCTGGTGCAGCAATGCAGAGATTTTGGTTCAACCTGTTCCAGAGAGATGGAT 60
DB 1 ATGGCCCTCCCTGGTGCAGCAATGCAGAGATTTTGGTTCAACCTGTTCCAGAGAGATGGAT 60
QY 61 GACAAATCAAGGAAATGGAATGTGTTCTTTCCCTCTGAGCCCTCTTGGCTGCCCTGGCC 120
DB 61 GACAAATCAAGGAAATGGAATGTGTTCTTTCCCTCTGAGCCCTCTTGGCTGCCCTGGCC 120
QY 121 CTGCTCGCTGGTGGCGCTCAGATGACTCCCTCTCAGATTTGATTAAGTTGTTTCATGTT 180
DB 121 CTGCTCGCTGGTGGCGCTCAGATGACTCCCTCTCAGATTTGATTAAGTTGTTTCATGTT 180
QY 181 AACACTGCTCAGATATGGAATCTTTCTTAATAGTACAGGCTCAGGCTCAGGTCCTCACTG 240
DB 181 AACACTGCTCAGATATGGAATCTTTCTTAATAGTACAGGCTCAGGCTCAGGTCCTCACTG 240
QY 241 AAAAGAGTTTCTCATATAAATGCTCCCAAGGATATGATCTGAGCTGCTGCTGCTGCTGCT 300
DB 241 AAAAGAGTTTCTCATATAAATGCTCCCAAGGATATGATCTGAGCTGCTGCTGCTGCTGCT 300
QY 301 GGGCTTTTCTCAAAAGCTATGCTGTTTCAATAGTACAGGCTCAGGCTCAGGTCCTCACTG 360
DB 301 GGGCTTTTCTCAAAAGCTATGCTGTTTCAATAGTACAGGCTCAGGCTCAGGTCCTCACTG 360
QY 361 TTATACGATGCCAAAGTGGAGCGAGTTGACTTTTACGAATCAATTTAGAGACACTAGAGCT 420
DB 361 TTATACGATGCCAAAGTGGAGCGAGTTGACTTTTACGAATCAATTTAGAGACACTAGAGCT 420
QY 421 AATATAATAGTGGGTTGAAATGAACACATGGCAAAATCAAGACGTTGATTTGTTGAA 480
DB 421 AATATAATAGTGGGTTGAAATGAACACATGGCAAAATCAAGACGTTGATTTGTTGAA 480
QY 481 GGTGGCATAGCTCATCTGCTGTAATGGTGTGCTGTAATGCTGTGCTGCTGCTGCTGCTGCT 540
DB 481 GGTGGCATAGCTCATCTGCTGTAATGGTGTGCTGTAATGCTGTGCTGCTGCTGCTGCTGCT 540
QY 541 TGGCATCAGCTTTCAACAGAGCGAAACCAATAAATGGCATTTCAATCTCCCAAGTGC 600
DB 541 TGGCATCAGCTTTCAACAGAGCGAAACCAATAAATGGCATTTCAATCTCCCAAGTGC 600
QY 601 TGTGGAGAGAGTGGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 601 TGTGGAGAGAGTGGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 GACCCATCAATGAGATTTCTGAGCTCAGATACATGCTGGCATTAACATGATGATGATGATGAT 720
DB 661 GACCCATCAATGAGATTTCTGAGCTCAGATACATGCTGGCATTAACATGATGATGATGATGAT 720
QY 721 CTGCTGAGATGACTCTCTCTGAAATTTGAAACAACTGACCTTTCAAGATCAATAGGAA 780

DB 721 CTGCTGAGATGACTCTCTCTGAAATTTGAAACAACTGACCTTTCAAGATCAATAGGAA 780
QY 781 TGGCAATCCCAAGGCGAATGACCTCTAAGTATGTTGAGGATTTTCTCTCAGTTCAAG 840
DB 781 TGGCAATCCCAAGGCGAATGACCTCTAAGTATGTTGAGGATTTTCTCTCAGTTCAAG 840
QY 841 ATAGAGAGATTTATGAATGAACAAATATTTGAGAGCCCTAGGCTGAAAGATATCTTT 900
DB 841 ATAGAGAGATTTATGAATGAACAAATATTTGAGAGCCCTAGGCTGAAAGATATCTTT 900
QY 901 GATGAATCCCAAGGCGAATGACCTCTCTGGGATTTCTCGGGGGGCTGCTCTATATCAAGG 960
DB 901 GATGAATCCCAAGGCGAATGACCTCTCTGGGATTTCTCGGGGGGCTGCTCTATATCAAGG 960
QY 961 ATGATGCACAAAATCTTACATAGAGTCTACTGAGAGGCGACCGAGGCTACTGCTGCACA 1020
DB 961 ATGATGCACAAAATCTTACATAGAGTCTACTGAGAGGCGACCGAGGCTACTGCTGCACA 1020
QY 1021 GGAAGTAAATTTGTAAGAAAGCAACTCCCTCAGTCCACGCTGTTTAGAGCTGACACCCA 1080
DB 1021 GGAAGTAAATTTGTAAGAAAGCAACTCCCTCAGTCCACGCTGTTTAGAGCTGACACCCA 1080
QY 1081 TTCTCTATTTGTTATCAGGAGAGTGAATCATCTTATTCAGTGGCAAGTTTCTTGCCT 1140
DB 1081 TTCTCTATTTGTTATCAGGAGAGTGAATCATCTTATTCAGTGGCAAGTTTCTTGCCT 1140
QY 1141 TGA 1143
DB 1141 TGA 1143

RESULT 2
AR052445 1950 bp DNA linear PAT 29-SEP-1999
LOCUS Sequence 30 from patent US 5831030.
DEFINITION
ACCESSION AR052445
VERSION AR052445.1 GI:5975809
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1950)
AUTHORS Teiimoto, M., Iwasa, F., Teiuroka, N., Nakazato, H., Miura, K.,
Ishida, N., Kuribara, T., Yamachi, K. and Yamaguchi, N.
TITLE Antibodies specific for megakaryocyte differentiation factor
JOURNAL Patent: US 5831030-A 30 03-NOV-1998;
FEATURES Location/Qualifiers
source 1..1950
/organism="unknown"

BASE COUNT 570 a 384 c 407 g 589 t
ORIGIN

Query Match 100.0%; Score 1143; DB 6; Length 1950;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTCCCTGGTGGCGCTCAGCAATGAGGATTTGCTTCAACCTGTTCCAGAGAGATGGAT 60
DB 74 ATGGCTCCCTGGTGGCGCTCAGCAATGAGGATTTGCTTCAACCTGTTCCAGAGAGATGGAT 133
QY 61 GACAAATCAAGGAAATGGAATGTGTTCTTTCCCTCTGAGCCCTCTTGGCTGCCCTGGCC 120
DB 134 GACAAATCAAGGAAATGGAATGTGTTCTTTCCCTCTGAGCCCTCTTGGCTGCCCTGGCC 193
QY 121 CTGCTCGCTGGCGCTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
DB 134 CTGCTCGCTGGCGCTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 253
QY 181 AACACTGCTCAGGATATGGAATCTTCTTAATAGTACAGGCTCCAGCTCTCAACTG 240
DB 254 AACACTGCTCAGGATATGGAATCTTCTTAATAGTACAGGCTCCAGCTCTCAACTG 313

241 AAAAGAGTTTTTCTGATATATAATGATCCCAAGAGATTATGATCTCAGCATGTGAAAT 300
 314 AAAAGAGTTTTTCTGATATATAATGATCCCAAGAGATTATGATCTCAGCATGTGAAAT 373
 301 GGGCTTTTCTGAAAAAGCTGATAGCTTTCTAAGAGCTACATTTAGTGTGCGGAAAAA 360
 374 GGGCTTTTCTGAAAAAGCTGATAGCTTTCTAAGAGCTACATTTAGTGTGCGGAAAAA 433
 361 TTATACGATCCAAAGTGGAGCGAGTTGACTTTTACGAATCATTTAGAGACACTAGACGT 420
 434 TTATACGATCCAAAGTGGAGCGAGTTGACTTTTACGAATCATTTAGAGACACTAGACGT 493
 421 AATATTAAATAGTGGTTGAAAAATGAAAAACATGTCGCAAAATCAAGAACGTGATGTGAA 480
 494 AATATTAAATAGTGGTTGAAAAATGAAAAACATGTCGCAAAATCAAGAACGTGATGTGAA 553
 481 GGTGGCATAGCTCATCTCTGTAATGCTGCTGTAATGCTGCTGTAATGCTGCTGTAATGCT 540
 554 GGTGGCATAGCTCATCTCTGTAATGCTGCTGTAATGCTGCTGTAATGCTGCTGTAATGCT 613
 541 TGGCAATCAGCCTTTCACCAAGAGCGAAACCAATAAATGCCATTTCAAATCTCCCAAGTGC 600
 614 TGGCAATCAGCCTTTCACCAAGAGCGAAACCAATAAATGCCATTTCAAATCTCCCAAGTGC 673
 601 TGTGGGAAGCAGTGGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 674 TGTGGGAAGCAGTGGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 733
 661 GACCCATCATAGTGAATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 734 GACCCATCATAGTGAATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 793
 721 CTGCTGAGATGACCTCTCTGAAATGAAACAAACTGACCTTTGCAATCTAATGAA 780
 794 CTGCTGAGATGACCTCTCTGAAATGAAACAAACTGACCTTTGCAATCTAATGAA 853
 781 TGGACCAATCCAAAGCGAATGACCTCTGAAATGAAACAAACTGACCTTTGCAATCTAATGAA 840
 854 TGGACCAATCCAAAGCGAATGACCTCTGAAATGAAACAAACTGACCTTTGCAATCTAATGAA 913
 841 ATAGAGAGAAATATGAAATGAAACAAATTTGAGAGCGCTAGGCTGAAAGATATCTTT 900
 914 ATAGAGAGAAATATGAAATGAAACAAATTTGAGAGCGCTAGGCTGAAAGATATCTTT 973
 901 GATGAATCCAAAGCAGATCTCTGGGATTTGCTTGGGGGGTCTCTGATATATCAAG 960
 974 GATGAATCCAAAGCAGATCTCTGGGATTTGCTTGGGGGGTCTCTGATATATCAAG 1033
 961 ATGATGCAAAATCTTACATAGAGTCACTGAGGAGGCGACCGAGGCTACTGCTGCCACA 1020
 1034 ATGATGCAAAATCTTACATAGAGTCACTGAGGAGGCGACCGAGGCTACTGCTGCCACA 1093
 1021 GGAAGTAATATTTAGAAAAAGCAACTCCCTCAGTCCACGCTGTTTAGAGCTGACCAACCA 1080
 1094 GGAAGTAATATTTAGAAAAAGCAACTCCCTCAGTCCACGCTGTTTAGAGCTGACCAACCA 1153
 1081 TTCCATTTGTTATCAGGAGGATGACATCATCTTATCAGTGGCAAGTTTCTTCCCT 1140
 1154 TTCCATTTGTTATCAGGAGGATGACATCATCTTATCAGTGGCAAGTTTCTTCCCT 1213
 1141 TGA 1143
 1214 TGA 1216

RESULT 3
 AR082433
 LOCUS 1950 bp DNA linear PAT 31-AUG-2000
 DEFINITION Sequence 30 from patent US 5972886.
 ACCESSION AR082433
 VERSION AR082433.1 GI:10009159
 KEYWORDS Unknown.

ORGANISM Unknown.
 REFERENCE Unclassified.
 AUTHORS 1 (bases 1 to 1950)
 Tsubimoto, M., Iwasa, F., Tsuruoka, N., Nakazato, H., Miura, K.,
 Ishida, N., Kurihara, T., Yamauchi, K. and Yamaguchi, N.
 TITLE Megakaryocyte differentiation factor
 JOURNAL Patent: US 5972886-A 30 26-OCT-1999;
 FEATURES Location/Qualifiers
 source 1..1950
 /organism="unknown"
 BASE COUNT 570 a 384 c 407 g 589 t
 ORIGIN
 Query Match 100.0%; Score 1143; DB 6; Length 1950;
 Local Similarity 100.0%; Pred. No. 0;
 Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGGCTCCCTTCTGTCAGCAAAATGCGAGATTTTGGCTTCAACCTGTTCAGAGAGATGAT 60
 DB 74 ATGGCTCCCTTCTGTCAGCAAAATGCGAGATTTTGGCTTCAACCTGTTCAGAGAGATGAT 133
 QY 61 GCAATCAAGAAATGGAATGATGTTTCTCTCTGAGCTCTTCGCTGCCCTGGCC 120
 DB 134 GCAATCAAGAAATGGAATGATGTTTCTCTCTGAGCTCTTCGCTGCCCTGGCC 193
 QY 121 CTGCTCCCTTGGGCGCTCAGATGACTCCCTCTCTCAGATTGATAAGTTGCTTCAATGTT 180
 DB 194 CTGCTCCCTTGGGCGCTCAGATGACTCCCTCTCTCAGATTGATAAGTTGCTTCAATGTT 253
 QY 181 AACCTGCTCAGGATGATAAGAACTCTTTAAATAGTCACTCAGGCTCCAGTCTCAACTG 240
 DB 254 AACCTGCTCAGGATGATAAGAACTCTTTAAATAGTCACTCAGGCTCCAGTCTCAACTG 313
 QY 241 AAAAGATTTTCTGATATAAATGCAATCCCAAGATTAATGATCTCAGCATGTGAAAT 300
 DB 314 AAAAGATTTTCTGATATAAATGCAATCCCAAGATTAATGATCTCAGCATGTGAAAT 373
 QY 301 GGGCTTTTCTGAAAAAGCTGATAGCTTTTCAATAAGAGCTACATTTAGTGTGCGGAAAAA 360
 DB 374 GGGCTTTTCTGAAAAAGCTGATAGCTTTTCAATAAGAGCTACATTTAGTGTGCGGAAAAA 433
 QY 361 TTATACGATCCAAAGTGGAGCGAGTTGACTTTACGAATCATTTAGAGACACTAGACGT 420
 DB 434 TTATACGATCCAAAGTGGAGCGAGTTGACTTTACGAATCATTTAGAGACACTAGACGT 493
 QY 421 AATATTAAATAGTGGTTGAAAAATGAAAAACATGTCGCAAAATCAAGAACGTGATGTGAA 480
 DB 494 AATATTAAATAGTGGTTGAAAAATGAAAAACATGTCGCAAAATCAAGAACGTGATGTGAA 553
 QY 481 GGTGGCATAGCTCATCTCTGTAATGCTGCTGTAATGCTGCTGTAATGCTGCTGTAATGCT 540
 DB 554 GGTGGCATAGCTCATCTCTGTAATGCTGCTGTAATGCTGCTGTAATGCTGCTGTAATGCT 613
 QY 541 TGGCAATCAGCCTTTCACCAAGAGCGAAACCAATAAATGCCATTTCAAATCTCCCAAGTGC 600
 DB 614 TGGCAATCAGCCTTTCACCAAGAGCGAAACCAATAAATGCCATTTCAAATCTCCCAAGTGC 673
 QY 601 TGTGGGAAGCAGTGGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 DB 674 TGTGGGAAGCAGTGGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 733
 QY 661 GACCCATCATAGTGAATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 DB 734 GACCCATCATAGTGAATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 793
 QY 721 CTGCTGAGATGACCTCTCTGAAATGAAACAAACTGACCTTTGCAATCTAATGAA 780
 DB 794 CTGCTGAGATGACCTCTCTGAAATGAAACAAACTGACCTTTGCAATCTAATGAA 853
 QY 781 TGGACCAATCCAAAGCGAATGACCTCTGAAATGAAACAAACTGACCTTTGCAATCTAATGAA 840
 DB 854 TGGACCAATCCAAAGCGAATGACCTCTGAAATGAAACAAACTGACCTTTGCAATCTAATGAA 913

961 ATGATGCAAAATCTTACATAGAGGTCACTGAGAGGGCAACGAGGCTACTGTCGCCACA 1020
1034 ATGATGCAAAATCTTACATAGAGGTCACTGAGAGGGCAACGAGGCTACTGTCGCCACA 1093
1021 GGAAGTAATATGTAGAAAGCAATCCCTCAGTCACTGAGGCTGTTAGAGCTGACCAACCA 1080
1094 GGAAGTAATATGTAGAAAGCAATCCCTCAGTCACTGAGGCTGTTAGAGCTGACCAACCA 1153
1081 TTCCTATTTGTTATCAGAGGATGACATCATCTTATTCACTGGCAAGTTTCTTCCCT 1140
1154 TTCCTATTTGTTATCAGAGGATGACATCATCTTATTCACTGGCAAGTTTCTTCCCT 1213
1141 TGA 1143
1214 TGA 1216

RESULT 5
D88575 1950 bp mRNA linear PRI 04-JUL-1998
LOCUS Homo sapiens mRNA for TP55, complete cds.
DEFINITION D88575
ACCESSION D88575.1 GI:3288674
VERSION TP55.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (sites)
Tsujimoto, M., Tsuruoka, N., Ishida, N., Kurihara, T., Iwasa, P.,
Yanashiro, K., Rogi, T., Kodama, S., Katsuragi, N., Adachi, M.,
Kakayama, T., Nakao, M., Yamauchi, K., Hashino, J., Haruyama, M.,
Miura, K., Nakanishi, T., Nakazato, H., Teramura, M., Mizoguchi, H. and
Yanaguchi, N.
Purification, cDNA cloning, and characterization of a new serpin
with megakaryocyte maturation activity
J. Biol. Chem. 272 (24), 15373-15380 (1997)
MEDLINE 97326116
PUBMED 9182567
2 (bases 1 to 1950)
Tsuruoka, N.
Direct Submission
Submitted (25-OCT-1996) Nobuo Tsuruoka, Suntory Institute for
Biomedical Research, Pharmaceutical Research Laboratories I; 1-1-1
Wakayamadai, Shinamoto-cho, Mishima-gun, Osaka 618, Japan
(E-mail:tsuruoka@minase.suntory.co.jp, Tel:075-962-9285(ex.3110),
Fax:075-962-6448)
Location/Qualifiers
1. 1950
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="A431"
/cell_type="epidermoid carcinoma"
74. 1216
/codon_start=1
/product="tp55"
/protein_id="BAA31232.1"
/db_xref="GI:3288675"

CDS
translation="MASLAAANAFPCNLFREMDNCGNVPFSSLSLPAALVRL
GAQDLSIQIDKLHVNTASGNSQSGHQQLKRVFSDINASHADYDLSYNGL
FAKYVGFHVDIECAEKLDAKVERVFTNHLEDTRNKWENETHGKIKRVIGE
GGTSSAVMVLVNAVYFKGWQSAFTKSETTINCFKSPKSGKAVAMMHQERKFLSV
LEDSKWLLELYNGINMYVLPENDLSEINLKFONLMEWTNPRMTSKYVEVFP
PQFKIKENTEMKQYRALGKIDIPESKADLSGIASGRLYISRMWHSYIEVTEGT
EATAATGSNIVEKQFPOSTFLRADHPFLFVIRKDDILFLSGRVSCP"

BASE COUNT 570 a 384 c 407 g 589 t
ORIGIN
Query Match 100.0%; Score 1143; DB 9; Length 1950;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTCCCTTGTCTGCGCAAAATGCAGAGATTTGTTCAACCTGTTCAAGAGAGATGGAT 60
DB 74 ATGGCTCCCTTGTCTGCGCAAAATGCAGAGATTTGTTCAACCTGTTCAAGAGAGATGGAT 133
QY 61 GACAAATCAAGGAAATGGAAATGTGTCTCTCTGAGGCTCTTCGCTGCGCTGGGC 120
DB 134 GACAAATCAAGGAAATGGAAATGTGTCTCTCTGAGGCTCTTCGCTGCGCTGGGC 193
QY 121 CTGGTCCGCTTGGGGCGCTCAAGATGATCTCCCTCTCTCAGATTGATAAGTTCGTTCACT 180
DB 194 CTGGTCCGCTTGGGGCGCTCAAGATGATCTCCCTCTCTCAGATTGATAAGTTCGTTCACT 253
QY 181 AACACTGCTCAGGATATGGAATCTCTCTAATAGTCACTGAGGCTCCACTCTCAACTG 240
DB 254 AACACTGCTCAGGATATGGAATCTCTCTAATAGTCACTGAGGCTCCACTCTCAACTG 313
QY 241 AAAAGAGTCTTTCTGATATAAATGCATCCCAAGGATTATGATCTCAGCAATTTGCAAT 300
DB 314 AAAAGAGTCTTTCTGATATAAATGCATCCCAAGGATTATGATCTCAGCAATTTGCAAT 373
QY 301 GGGCTTTTCTGAAAAAGTGTATGGCTTTCAATAGGACTACATTCAGTGTGCCGAAAAA 360
DB 374 GGGCTTTTCTGAAAAAGTGTATGGCTTTCAATAGGACTACATTCAGTGTGCCGAAAAA 433
QY 361 TTATACGATGCCAAAGTGGAGGTTGACTTTACGAATCAATTTAGAAACACACTAGACGT 420
DB 434 TTATACGATGCCAAAGTGGAGGTTGACTTTACGAATCAATTTAGAAACACACTAGACGT 493
QY 421 AATATTAAATAGTGGGTTGAAAAATGAAACACATGCAAAATCAAGAACCTGATTTGGTAA 480
DB 494 AATATTAAATAGTGGGTTGAAAAATGAAACACATGCAAAATCAAGAACCTGATTTGGTAA 553
QY 481 GGTGGCAATAGCTCACTGCTGTTAAAGTGTGTTGAATGCTGTGACTTCAAGGCGAG 540
DB 554 GGTGGCAATAGCTCACTGCTGTTAAAGTGTGTTGAATGCTGTGACTTCAAGGCGAG 613
QY 541 TGGCAATTCAGCTTCCACAGAGCGAAACCAATAATTCGCAATTTCAAAATCTCCCAAGTGC 600
DB 614 TGGCAATTCAGCTTCCACAGAGCGAAACCAATAATTCGCAATTTCAAAATCTCCCAAGTGC 673
QY 601 TCTGGAGGCGAGTCCCATGATCATCAGGAACGGAAGTTCAATTTGCTGTTATTGAG 660
DB 674 TCTGGAGGCGAGTCCCATGATCATCAGGAACGGAAGTTCAATTTGCTGTTATTGAG 733
QY 661 GACCAATCAATGAAAGATTCTTGACTCAGTACATGCTGGCTAAACATGTAACATGTAAC 720
DB 734 GACCAATCAATGAAAGATTCTTGACTCAGTACATGCTGGCTAAACATGTAACATGTAAC 793
QY 721 CTGGCTGAGATGACCTCTCTGAAATTTGAAACCAAACTGACCTTTCAGAAATCTAATGGAA 780
DB 794 CTGGCTGAGATGACCTCTCTGAAATTTGAAACCAAACTGACCTTTCAGAAATCTAATGGAA 853
QY 781 TGGACCAATCCAAAGGCGAATGCTCTAAGTATGTTAGGATTTTTTTCCTCAGTTCAAG 840
DB 854 TGGACCAATCCAAAGGCGAATGCTCTAAGTATGTTAGGATTTTTTTCCTCAGTTCAAG 913
QY 841 ATAGAGAAATATTAATGAAATGAAACATATTTTCAGAGCCCTAGGCTGAAAGATATCTTT 900
DB 914 ATAGAGAAATATTAATGAAATGAAACATATTTTCAGAGCCCTAGGCTGAAAGATATCTTT 973
QY 901 GATGAATCCAAAGCAGATCTCTCTGGGATTTGCTTCGGGGGGTTCGTTGTATATATCAAG 960
DB 974 GATGAATCCAAAGCAGATCTCTCTGGGATTTGCTTCGGGGGGTTCGTTGTATATATCAAG 1033
QY 961 ATGATGCAAAATCTTACATAGAGTCACTGAGAGGGCACCGAGGCTACTGCTGCCACA 1020
DB 1034 ATGATGCAAAATCTTACATAGAGTCACTGAGAGGGCACCGAGGCTACTGCTGCCACA 1093
QY 1021 GGAAGTAATATGTAGAAAGCAATCCCTCAGTCACTGAGGCTGTTAGAGCTGACCAACCA 1080
DB 1094 GGAAGTAATATGTAGAAAGCAATCCCTCAGTCACTGAGGCTGTTAGAGCTGACCAACCA 1153

Mon Dec 15 08:51:16 2003

```

1081 TTCCTATTGTTATCAGGAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGCCCT 1140
Db      TTCCCTATTTGTTATCAGGAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGCCCT 1213
1154 TTCCCTATTTGTTATCAGGAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGCCCT 1213
QY      1141 TGA 1143
Db      1214 TGA 1216

RESULT 6
AF027866      2249 bp mRNA linear PRI 20-OCT-1998
LOCUS      Homo sapiens megin mRNA, complete cds.
DEFINITION
ACCESSION      AF027866
VERSION
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM
REFERENCE
AUTHORS      Miyata, T., Nangaku, M., Suzuki, D., Inagi, R., Uragami, K., Sakai, H.,
Okubo, K. and Kurokawa, K.
TITLE      A mesangium-predominant gene, megin, is a new serpin upregulated
in IGA nephropathy
JOURNAL      J. Clin. Invest. 102 (4), 828-836 (1998)
MEDLINE      98376492
PUBMED      9710452
REFERENCE
AUTHORS      Toshio, M., Yasuda, Y., Nangaku, M. and Kurokawa, K.
TITLE      Direct Submission
JOURNAL      Submitted (02-OCT-1997) Medical Science, Tokai University,
Bousaidai, Isehara, Kanagawa 259-11, Japan
REFERENCE
AUTHORS      Toshio, M., Yasuda, Y., Nangaku, M. and Kurokawa, K.
TITLE      Direct Submission
JOURNAL      Submitted (19-OCT-1998) Medical Science, Tokai University,
Bousaidai, Isehara, Kanagawa 259-11, Japan
REMARK      Sequence update by submitter
COMMENT      On Oct 20, 1998 this sequence version replaced gi:3764058.
FEATURES
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_type="mesangial"
365..1507
/codon_start=1
/product="megin"
/protein_id="AAC64506.1"
/db_xref="GI:3769373"
/translation="MASLAAAEAEFCNLPRMDNCGNVTSSLSLFAALALVRL
GADQLSLQIDKLHVNATAGYNSNSQSLQSLKRVSPDINASHKDYDLSIVNGL
FASKVYGFKYDIEAEKLYDAKRVDFTHLDETRNINWENTHKGKINWIGE
GGTSSAAMVLVNAVYFKGWQSAFTKSETINLHFKSPKSGKAVAMVHQKRNLSV
IEDPSMKLLELVNGGIMNVILPNDLSIEINKLAFQNLWENTPRMTSKVEVFF
PQKIEKNEMKQVIRALGLKDI FDSKADLSGIASGRLYI SRMHKSYIEVTEBGT
EATAEGSNIVEKLPQSTLFRADHPFLFVIRKDDILFSGKVSQP"
BASE COUNT      649 a 452 c 461 g 583 t 4 others
ORIGIN
Query Match      100.0%; Score 1143; DB 9; Length 2249;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 ATGGCTCCCTTGCTGCAGCAATGCAGATTGCTTCAACCTGTTTCAGAGATGAT 60
Db      365 ATGGCTCCCTTGCTGCAGCAATGCAGATTGCTTCAACCTGTTTCAGAGATGAT 424
QY      61 GACAATCAGGAATGGAATGTGTTCTTCTCTCTGAGCTCTCGGTCGCCGTCGCC 120
Db      425 GACAATCAGGAATGGAATGTGTTCTTCTCTCTGAGCTCTCGGTCGCCGTCGCC 484
121 CTGGTCCGCTTGCGGCTCAAGTGAATCTCCCTCTCTCAGATTGATAAGTTGCTTCATGTT 180
485 CTGGTCCGCTTGCGGCTCAAGTGAATCTCCCTCTCTCAGATTGATAAGTTGCTTCATGTT 544
181 AACACTGGCTCAGATATGGAACCTCTTCTAATAGTCACTCAGGGCTCCAGTCTCAACTG 240
545 AACACTGGCTCAGGATATGGAACCTCTTCTAATAGTCACTCAGGGCTCCAGTCTCAACTG 604
241 AAAGAGTTTTTCTGATATAATGCATCCCAAGGATTAATGATCTCAGCATGTTGAAT 300
605 AAAGAGTTTTTCTGATATAATGCATCCCAAGGATTAATGATCTCAGCATGTTGAAT 664
301 GGGCTTTTCTGAAAAAGTTGCTTCTAATAGGATACATTTAGTGTGCGGAAAAA 360
665 GGGCTTTTCTGAAAAAGTTGCTTCTAATAGGATACATTTAGTGTGCGGAAAAA 724
361 TTATACGATGCCAAGTGGAGCGAGTTGACTTTAGCAATCAATTTAGAAGACACTAGCGT 420
725 TTATACGATGCCAAGTGGAGCGAGTTGACTTTAGCAATCAATTTAGAAGACACTAGCGT 784
421 AATATTAATAGTGGGTTGAAAAATGAAAAACAATGGCAAAATCAAGAACGTGATTGGTAA 480
785 AATATTAATAGTGGGTTGAAAAATGAAAAACAATGGCAAAATCAAGAACGTGATTGGTAA 844
481 GGTGGCATAAGCTCATCTGCTGTAATGCTGCTGTAATGCTGTAATGCTGTAATGCTG 540
845 GGTGGCATAAGCTCATCTGCTGTAATGCTGCTGTAATGCTGTAATGCTGTAATGCTG 904
541 TGGCAATCAGGCTTCACCAAGAGGCAAAACATAAATGGCAATTTCAAAATCTCCAGTGC 600
905 TGGCAATCAGGCTTCACCAAGAGGCAAAACATAAATGGCAATTTCAAAATCTCCAGTGC 964
601 TCTGGCAAGGCAATGGCCATGATGATCAGGAACGGAAGTTCAATTTGCTGTTATGAG 660
965 TCTGGCAAGGCAATGGCCATGATGATCAGGAACGGAAGTTCAATTTGCTGTTATGAG 1024
661 GACCCATCAATGAAGATTCTTGAGCTCAGATACATAATGGTGGCATAAACATGACGTTG 720
1025 GACCCATCAATGAAGATTCTTGAGCTCAGATACATAATGGTGGCATAAACATGACGTTG 1084
721 CTGGCTGGAATGACCTCTCTGAAATGAAAAAACAACGACCTTTCAAGATTAATGGAA 780
1085 CTGGCTGGAATGACCTCTCTGAAATGAAAAAACAACGACCTTTCAAGATTAATGGAA 1144
781 TGGACCAATCCAAAGGCGAATGACCTCTAAGTATGTTGAGGTATTTTCTTCAGTCAAG 840
1145 TGGACCAATCCAAAGGCGAATGACCTCTAAGTATGTTGAGGTATTTTCTTCAGTCAAG 1204
841 ATAGAGAGAAATTAATGAATGAACAAATATTTGAGAGCCCTAGGCTGGAAGATATCTTT 900
1205 ATAGAGAGAAATTAATGAATGAACAAATATTTGAGAGCCCTAGGCTGGAAGATATCTTT 1264
901 GATGATCCAAAGCAGATCTCTCTGGATGTTCTCGGGGCTGCTCTGTATATATCAAGG 960
1265 GATGATCCAAAGCAGATCTCTCTGGATGTTCTCGGGGCTGCTCTGTATATATCAAGG 1324
961 ATGATGCACAAATCTTATATAGAGTCACTGAGGAGGACCCAGGCTACTGTGCCACA 1020
1325 ATGATGCACAAATCTTATATAGAGTCACTGAGGAGGACCCAGGCTACTGTGCCACA 1384
1021 GGAATTAATTTATAGAAAGCAATCCCTCAGTCCAGCTGTTTATAGCTGACCCCA 1080
1385 GGAATTAATTTATAGAAAGCAATCCCTCAGTCCAGCTGTTTATAGCTGACCCCA 1444
1081 TTCCCTATTTGTTATCAGGAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGCCCT 1140
1445 TTCCCTATTTGTTATCAGGAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGCCCT 1504
1141 TGA 1143
1505 TGA 1507

```

Mon Dec 15 08:51:16 2003

RESULT 7

AC072051/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

AC072051 157284 bp DNA linear PRI 03-DEC-2001

Homo sapiens chromosome, clone RP11-79D21, complete sequence.

AC072051.8 GI:16974280

RTG.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 157284)

Birren, B., Linton, L., Nusbaum, C. and Lander, E.

Homo sapiens chromosome, clone RP11-79D21

2 (bases 1 to 157284)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,

Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,

Boguslavsky, L., Bouckhgalter, B., Brown, A., Burkett, G.,

Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,

Collimore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,

Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,

Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,

Grand-Pierre, N., Grant, G., Hagos, B., Hearford, A., Horton, L.,

Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kam, L., Karatas, A.,

Klein, J., Lakoque, K., Lamazares, R., Landers, T., Lehoczy, J.,

Levine, R., Liu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,

McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,

Meldrim, J., Menes, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,

Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Neill, D.,

O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,

Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,

Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,

Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,

Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,

Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,

Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission

Submitted (07-JUN-2000) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 157284)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,

Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouckhgalter, B.,

Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,

Choepel, Y., Colangelo, M., Collins, S., Collimore, A., Cooke, A.,

Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,

Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,

Ginde, S., Goid, S., Goyette, M., Graham, L., Grand-Pierre, N.,

Hagos, B., Hearford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,

Jones, C., Kamat, A., Karatas, A., Kells, C., Lakoque, K.,

Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,

MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,

McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J.,

Menes, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,

Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,

Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,

Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,

Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,

Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,

Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,

Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,

Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,

Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (25-NOV-2001) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 157284)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,

Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouckhgalter, B.,

Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,

Choepel, Y., Colangelo, M., Collins, S., Collimore, A., Cooke, A.,

Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,

Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,

Ginde, S., Goid, S., Goyette, M., Graham, L., Grand-Pierre, N.,

Hagos, B., Hearford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,

Jones, C., Kamat, A., Karatas, A., Kells, C., Lakoque, K.,

Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,

MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,

McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J.,

Menes, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,

Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,

Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,

Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,

Ginde, S., Goid, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Hearford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Kamat, A., Karatas, A., Kells, C., Lakoque, K.,
Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J.,
Menes, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (03-DEC-2001) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Nov 18, 2001 this sequence version replaced gi:14277307.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: MIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L366

Center Clone name: 79_D_21

Location/Qualifiers

1. 157284

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/clone="RP11-79D21"

/clone_lib="RP11-79D21 Human Male BAC"

599..635

/rpt_family="AT rich"

1040..1077

/rpt_family="AT rich"

1411..1747

/rpt_family="L1P"

2832..2992

/rpt_family="MLT1A1"

3115..3138

/rpt_family="AT rich"

complement(3487..3828)

/rpt_family="L1MC4a"

3834..4338

/rpt_family="MLT1C"

complement(4339..5397)

/rpt_family="L1MC4a"

complement(5742..5899)

/rpt_family="MER5B"

6499..6816

/rpt_family="HAL1"

complement(7416..7621)

/rpt_family="MER3"

complement(8163..9090)

/rpt_family="L2"

9276..9419

/rpt_family="MIR"

9598..9658

/rpt_family="MER81"

9659..10033

/rpt_family="MLT1J"

13986..14009

/rpt_family="AT rich"

14011..14115

/rpt_family="TA)n"

FEATURES

source

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

QY	925	GGGATTGCTTCGGGGGTCGTCTCTATATATCAAGAGTATGACAAATCTTATCATAGAG	984
Db	125931	GGGATTGCTTCGGGGGTCGTCTCTATATATCAAGAGTATGACAAATCTTATCATAGAG	125872
QY	985	GTCACTGAGGAGGACCCAGAGCTACTGCTGCCACAGAGAGTAATATTCAGAAAACCAA	1044
Db	125871	GTCACTGAGGAGGACCCAGAGCTACTGCTGCCACAGAGAGTAATATTCAGAAAACCAA	125812
QY	1045	CTCCCTCAGTCCACGCTGTGTAGAGCTGACCCACCATTCCTATTTCCTATTCAGGAGGAT	1104
Db	125811	CTCCCTCAGTCCACGCTGTGTAGAGCTGACCCACCATTCCTATTTCCTATTCAGGAGGAT	125752
QY	1105	GACATCATCTTATTCAGTGGCAAGTTCCTTGCCCTTGA	1143
Db	125751	GACATCATCTTATTCAGTGGCAAGTTCCTTGCCCTTGA	125713

RESULT	8
AC027524	linear
LOCUS	HTG 27-MAR-2003
DEFINITION	Homo sapiens chromosome 18 clone RP11-317G1 map 18, 3 unordered pieces.
ACCESSION	AC027524
KEYWORDS	HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_CANCELLED.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 189092)
TITLE	Homo sapiens chromosome 18, clone RP11-317G1
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 189092)
AUTHORS	Birren, B., Linton, J., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F., Boguslavsky, I., Bouckgalter, B., Brown, A., Burkett, G., Campolano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J. S., Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., LaRoque, K., Lamazares, K., Landers, T., Lechokky, C., Levine, R., Lieu, C., Liu, G., Locke, K., MacDonald, P., Marquis, N., McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R., McDermid, J., Menues, L., Mihova, T., Miranda, C., Mieng, V., Morrow, J., Murphy, T., Naylor, J., Norman, C. H., Peterson, K., Pier, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliou, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.
TITLE	Direct Submission
JOURNAL	Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE	3 (bases 1 to 189092)
AUTHORS	Birren, B., Linton, J., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, I., Bouckgalter, B., Brown, A., Camarata, J., Campolano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J. S., Dodge, S., Cook, A., Cooke, P., DeArelano, K., FitzHugh, W., Gage, D., Faro, S., Ferreira, P., Fitzgerald, M., FitzHugh, W., Goyette, M., Graham, L., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Hulme, W., Iliev, I., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Kells, C., LaRoque, K., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kellis, R., Lindblad-Toh, K., Lamazares, R., Landers, T., Lechokky, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McDermid, J., Menues, L., Mihova, T., Mieng, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., North, C., Norman, C. H., O'Connor, T., O'Donnell, P.,

[illegible]


```

Query Match      20.1%; Score 230; DB 6; Length 316;
Best Local Similarity 99.6%; Pred. No. 2.9e-120; Indels 0; Gaps 0;
Matches 280; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 822 ATTTTTCCTCAGTCAAGATAGAGAGAAATATGAATCAAAACAATATTGAGAGCCCT 881
Db 1 ATTTTTCCTCAGTCAAGATAGAGAGAAATATGAATCAAAACAATATTGAGAGCCCT 60

QY 882 AGGCTGAAGATATCTTTGATGATCCAAAGCAGATCTCTCTGGATTGCTCGGGGG 941
Db 61 AGGCTGAAGATATCTTTGATGATCCAAAGCAGATCTCTCTGGATTGCTCGGGGG 120

QY 942 TCCTCTGATATATCAGAGTATGATGATGATGATGATGATGATGATGATGATGAT 1001
Db 121 TCCTCTGATATATCAGAGTATGATGATGATGATGATGATGATGATGATGATGAT 180

QY 1002 CGAGGCTACTGCTGCCAGGAAAGTAAATATTGTAGAAAGCAACTCCCTCAGTCCAGCT 1061
Db 181 CGAGGCTACTGCTGCCAGGAAAGTAAATATTGTAGAAAGCAACTCCCTCAGTCCAGCT 240

QY 1062 GTTTAGAGTGCACCCATTCCTATTTCTTATCAGGAGG 1102
Db 241 GTTTAGAGTGCACCCATTCCTATTTCTTATCAGGAGG 281

RESULT 11
AR272521/c
LOCUS      AX062606
DEFINITION Sequence 233 from patent US 6504010.
ACCESSION  AR272521
VERSION     AR272521.1 GI:29704406
KEYWORDS   Unknown.
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedvick, T.S.,
AUTHORS     Carter, D., Retter, M.W., Mannion, J., and Fan, L.
TITLE       Compositions and methods for the therapy and diagnosis of lung
JOURNAL     Patent: US 6504010-A 233 07-JAN-2003;
FEATURES    Location/Qualifiers
source      1..600
            /organism="unknown"
BASE COUNT  198 a 106 c 135 g 161 t
ORIGIN

Query Match      10.6%; Score 121; DB 6; Length 600;
Best Local Similarity 100.0%; Pred. No. 1.8e-57;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1023 AAGTAATATTGTAGAAAAGCAACTCCCTCAGTCCAGCTGTTTAGAGCTGACCAACCCATT 1082
Db 600 AAGTAATATTGTAGAAAAGCAACTCCCTCAGTCCAGCTGTTTAGAGCTGACCAACCCATT 541

QY 1083 CCTATTGTTATCAGGAAGGATGACATCATCTTATTCAGTGGCAAAAGTTTCTTGCCTTG 1142
Db 540 CCTATTGTTATCAGGAAGGATGACATCATCTTATTCAGTGGCAAAAGTTTCTTGCCTTG 481

QY 1143 A 1143
Db 480 A 480

RESULT 13
AX062606/c
LOCUS      AX062606
DEFINITION Sequence 233 from Patent WO0100828.
ACCESSION  AX062606
VERSION     AX062606.1 GI:12540481
KEYWORDS   Homo sapiens (human)
SOURCE      Homo sapiens
ORGANISM    Homo sapiens
REFERENCE   Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedvick, T.S.,
AUTHORS     Carter, D., Retter, M.W. and Mannion, J.
TITLE       Compositions and methods for the therapy and diagnosis of lung
JOURNAL     Patent: WO 0100828-A 233 04-JAN-2001;
FEATURES    Location/Qualifiers
source      1..600
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
BASE COUNT  198 a 106 c 135 g 161 t
ORIGIN

Query Match      10.6%; Score 121; DB 6; Length 600;
Best Local Similarity 100.0%; Pred. No. 1.8e-57;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1023 AAGTAATATTGTAGAAAAGCAACTCCCTCAGTCCAGCTGTTTAGAGCTGACCAACCCATT 1082
Db 600 AAGTAATATTGTAGAAAAGCAACTCCCTCAGTCCAGCTGTTTAGAGCTGACCAACCCATT 541

QY 1083 CCTATTGTTATCAGGAAGGATGACATCATCTTATTCAGTGGCAAAAGTTTCTTGCCTTG 1142
Db 540 CCTATTGTTATCAGGAAGGATGACATCATCTTATTCAGTGGCAAAAGTTTCTTGCCTTG 481

QY 1143 A 1143
Db 480 A 480

RESULT 12
AR276102/c
LOCUS      AR276102
DEFINITION Sequence 233 from patent US 6509448.
ACCESSION  AR276102
VERSION     AR276102.1 GI:29709747
KEYWORDS   Unknown.
SOURCE      Unknown.

```


RESULT 14
AX367523/c
LOCUS AX367523 600 bp DNA linear PAT 16-FEB-2002
DEFINITION Sequence 233 from Patent WO0204514.
ACCESSION AX367523
VERSION AX367523.1 GI:18855627
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
Wang, T., Watanabe, Y., Henderson, R.A., Johnson, J.C., Retter, M.W.,
Marnerakis, M., Carter, D., Fanger, G.R., Vedvick, T.S., Bangur, C.S.,
McNabb, A., Fanger, N., Switzer, A., McNeill, P.D., and Clapper, J.D.
Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: WO 0204514-A 233 17-JAN-2002;
CORIXA CORPORATION (US)
FEATURES
Location/Qualifiers
source 1..600
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 198 a 106 c 135 g 161 t
ORIGIN
Query Match 10.6%; Score 121; DB 6; Length 600;
Best Local Similarity 100.0%; Pred. No. 1.8e-57; Indels 0; Gaps 0;
Matches 121; Conservative 0; Mismatches 0;
Qy 1023 AAGTAAATCTGAGAAAGCAACTCCCTCAGTCACCGCTGTGTTAGCTGACACCCATT 1082
Db 600 AAGTAAATCTGAGAAAGCAACTCCCTCAGTCACCGCTGTGTTAGCTGACACCCATT 541
Qy 1083 CCTATTCTTTCAGGAGAGATGACATCACTTATTCAGTGCACAAAGTTCTGCCCCTG 1142
Db 540 CCTATTCTTTCAGGAGAGATGACATCACTTATTCAGTGCACAAAGTTCTGCCCCTG 481
Qy 1143 A 1143
Db 480 A 480

RESULT 15
AF105329
LOCUS AF105329 1143 bp mRNA linear ROD 19-OCT-2001
DEFINITION Rattus norvegicus megin mRNA, complete cds.
ACCESSION AF105329
VERSION AF105329.1 GI:16269520
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 1143)
Nangaku, M., Miyata, T., Suzuki, D., Umezono, T., Hashimoto, T.,
Wada, T., Yagi, M., Nagano, N., Inagi, R. and Kurokawa, K.
Cloning of rodent megin revealed its up-regulation in
mesangio proliferative nephritis
JOURNAL Kidney Int. 60 (2), 641-652 (2001)
MEDLINE 21368006
PUBMED 11473647
REFERENCE 2 (bases 1 to 1143)
AUTHORS Miyata, T., Nangaku, T., Inagi, R. and Kurokawa, K.
TITLE Direct Submission
JOURNAL Submitted (09-NOV-1998) Institute of Medical Sciences, Tokai
University, Bohseidai, Isehara, Kanagawa 259-1193, Japan
FEATURES
Location/Qualifiers
source 1..1143
/organism="Rattus norvegicus"

Query Match 2.7%; Score 31; DB 10; Length 1143;
Best Local Similarity 100.0%; Pred. No. 1.3e-05; Indels 0; Gaps 0;
Matches 31; Conservative 0; Mismatches 0;
Qy 437 TTGAAATGAAACACATGGCAAAATCAAGAA 467
Db 437 TTGAAATGAAACACATGGCAAAATCAAGAA 467

Search completed: December 14, 2003, 06:01:42
Job time : 4377 secs

BASE COUNT 345 a 232 c 261 g 305 t
ORIGIN
Query Match 2.7%; Score 31; DB 10; Length 1143;
Best Local Similarity 100.0%; Pred. No. 1.3e-05; Indels 0; Gaps 0;
Matches 31; Conservative 0; Mismatches 0;
Qy 437 TTGAAATGAAACACATGGCAAAATCAAGAA 467
Db 437 TTGAAATGAAACACATGGCAAAATCAAGAA 467

Search completed: December 14, 2003, 06:01:42
Job time : 4377 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

ON nucleic - nucleic search, using sw model

Run on: December 14, 2003, 03:19:41 ; Search time 377 Seconds
(without alignments)
8184.238 Million cell updates/sec

Title: US-09-508-997A-1

Perfect score: 1143

Sequence: 1 atggctccctgtgcagc.....gcaagtttttgccttga 1143

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_15Jun03.*

```

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1143	100.0	1143	20	AA56712 Human megasin codin
2	1143	100.0	1143	21	AA99294 Human megasin nucle
3	1143	100.0	1143	22	AAF82438 Human megasin nucle
4	1143	100.0	1950	15	AAQ56670 Human megakaryocyt
5	1143	100.0	2249	25	ABX76260 Lung cancer-associ
6	254	22.2	254	25	ACA56060 Human signalling p
7	230	20.1	316	25	ACA56226 Human signalling p
8	121	10.6	600	22	AAF68315 Human lung tumour

C	9	121	10.6	600	24	ABK38226	cDNA encoding clon
C	10	121	10.6	600	25	ACA10555	Human lung cancer-
C	11	121	10.6	600	25	ABX99506	Lung cancer therap
C	12	31	2.7	42	20	AA56727	WO9915652 primer 1
	13	31	2.7	1147	20	AA56714	Mouse megasin DNA.
	14	31	2.7	1147	21	AA56714	Mouse megasin nucle
	15	31	2.7	1147	22	AAF82440	Mouse megasin nucle
	16	31	2.7	1229	20	AA56713	Rat megasin nucleot
	17	31	2.7	1229	21	AA56713	Rat megasin nucleot
	18	31	2.7	1229	22	AAH48181	Rat megasin protein
	19	31	2.7	1229	22	AAF82439	Rat megasin nucleot
	20	30	2.6	30	22	AAF82447	PCR primer hM8-1.
	21	30	2.6	44	20	AA56747	WO9915652 primer 3
	22	30	2.6	44	22	AAF82441	Human megasin prime
	23	29	2.5	38	20	AA56726	WO9915652 primer 1
C	24	26	2.3	26	22	AAF82442	Human megasin prime
	25	25	2.2	37	15	AAQ56668	Human megakaryocyt
	26	23	2.0	23	15	AAQ56651	Human megakaryocyt
	27	23	2.0	29	20	AA56722	WO9915652 primer 8
C	28	22	1.9	22	20	AA56750	WO9915652 primer 3
	29	22	1.9	29	20	AA56752	WO9915652 primer 3
C	30	21	1.8	21	15	AAQ56652	Human megakaryocyt
C	31	21	1.8	21	15	AAQ56653	Human megakaryocyt
C	32	21	1.8	21	15	AAQ56654	Human megakaryocyt
C	33	21	1.8	21	15	AAQ56655	Human megakaryocyt
C	34	21	1.8	21	15	AAQ56656	Human megakaryocyt
C	35	21	1.8	21	20	AA56749	WO9915652 primer 1
	36	21	1.8	21	20	AA56724	WO9915652 primer 3
C	37	21	1.8	27	20	AA56748	WO9915652 primer 3
C	38	21	1.8	30	20	AA56751	WO9915652 primer 9
C	39	21	1.8	30	20	AA56723	Human transporter
	40	21	1.8	34667	24	AA44328	Human transmembrin ge
	41	21	1.8	249999	25	ABZ80229	Human GDP-mannose
	42	20	1.7	343	25	ABX20549	B. burgdorferi ant
	43	20	1.7	1050	20	AA561536	B. burgdorferi ant
	44	20	1.7	1152	20	AA561595	Drosophila melanog
C	45	20	1.7	4493	23	ABL28683	

ALIGNMENTS

RESULT 1
AA56712
ID AAX56712 standard; DNA; 1143 BP.

XX AAX56712;

XX DT 14-JUN-1999 (first entry)

XX DE Human megasin coding region.

XX ME Megsin; mesangial cell; treatment; diagnosis; disease; IgA nephropathy;

XX KW human; rat; murine; ss.

XX OS Homo sapiens.

XX PN WO9915652-AL.

XX PD 01-APR-1999.

XX PF 22-SEP-1998; 98WO-JP04269.

XX PR 22-SEP-1997; 97JP-0275302.

XX (KUPO/) KUROKAWA K.

XX (MIYA/) MIYATA T.

XX PI Miyata T;

XX WPI; 1999-275983/23.

XX P-PSDB; AAY08254.

PT Megsin protein expressed specifically in mesangial cells

PS Claim 4; Page 58-61; 100pp; Japanese.

XX This invention describes the isolation of novel mesin nucleic acid and
 CC proteins from human, rat and mouse tissue. This protein is expressed
 CC specifically in mesangial cells. The products of the invention are
 CC useful for the treatment and diagnosis of diseases involving mesangial
 CC cells, such as IgA nephropathy.

XX SQ Sequence 1143 BP; 336 A; 230 C; 261 G; 316 T; 0 other;

Query Match 100.0%; Score 1143; DB 20; Length 1143;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCCCTCCCTGCTGAGCAATGCGAGTGTTCCTCAACCTGTCAGAGATGGAT 60
 Db 1 ATGGCCCTCCCTGCTGAGCAATGCGAGTGTTCCTCAACCTGTCAGAGATGGAT 60
 Qy 61 GACCAATCAAGCAATGGAATGCTCTCTTCTCTCTGAGCCCTTCGCTGCCCTGGCC 120
 Db 61 GACCAATCAAGCAATGGAATGCTCTCTTCTCTCTGAGCCCTTCGCTGCCCTGGCC 120
 Qy 121 CTGGTCCGCTTGGGCGCTCAAGATGACTCCCTCTCTGAGATGATAGTTCATGTT 180
 Db 121 CTGGTCCGCTTGGGCGCTCAAGATGACTCCCTCTCTGAGATGATAGTTCATGTT 180
 Qy 181 AACACTGCTCAGATATGGAATCTCTCTATAGTCTGAGTCCGCTCAGCTCAACTG 240
 Db 181 AACACTGCTCAGATATGGAATCTCTCTATAGTCTGAGTCCGCTCAGCTCAACTG 240
 Qy 241 AAAAGAGTTTTTCTGATATAATGCAATCCCAAGGATATGATCTCAGCATTTGCAAT 300
 Db 241 AAAAGAGTTTTTCTGATATAATGCAATCCCAAGGATATGATCTCAGCATTTGCAAT 300
 Qy 301 GGGCTTTTCTGAAAAGTGTATGGCTTTCTAAGGACTACATGAGTGTCCGAAAAA 360
 Db 301 GGGCTTTTCTGAAAAGTGTATGGCTTTCTAAGGACTACATGAGTGTCCGAAAAA 360
 Qy 361 TTATACAGTCCAAAGTGGAGGAGTTCAGTTCAGCATCTTTAGAGACTAGAGCT 420
 Db 361 TTATACAGTCCAAAGTGGAGGAGTTCAGTTCAGCATCTTTAGAGACTAGAGCT 420
 Qy 421 AATATTAAGTGGCTTGAATAATGAACACATGCGCAAAATCAAGAACTGATCGTAA 480
 Db 421 AATATTAAGTGGCTTGAATAATGAACACATGCGCAAAATCAAGAACTGATCGTAA 480
 Qy 481 GGTGGCAATAGCTCATCTCTCTAATGGTGTGGTGAATGCTGTGCTTCAAGGCAAG 540
 Db 481 GGTGGCAATAGCTCATCTCTAATGGTGTGGTGAATGCTGTGCTTCAAGGCAAG 540
 Qy 541 TGGCAATCAGCTTCCCAAGAGCGAATCAATGCAATTTCAATCTCCCAAGTGC 600
 Db 541 TGGCAATCAGCTTCCCAAGAGCGAATCAATGCAATTTCAATCTCCCAAGTGC 600
 Qy 601 TCTGGGAAGGAGCGGCAATGATGATCAGAAACGGAAGTTCATTTCTGTATTAAG 660
 Db 601 TCTGGGAAGGAGCGGCAATGATGATCAGAAACGGAAGTTCATTTCTGTATTAAG 660
 Qy 661 GACCCATCAATGAGATTTCTGAGCTCAGATACAAATGGTGGCAATGATGATGCTG 720
 Db 661 GACCCATCAATGAGATTTCTGAGCTCAGATACAAATGGTGGCAATGATGATGCTG 720
 Qy 721 CTGCTGAGATGACCTCTGGAATGGAACAACTGACCTTTCAAGATCTAATGGAA 780
 Db 721 CTGCTGAGATGACCTCTGGAATGGAACAACTGACCTTTCAAGATCTAATGGAA 780
 Qy 781 TGGACCAATCCAGGCGAATGACTCTAAGTATGATGATTTTTTCTCAGTTCAAG 840
 Db 781 TGGACCAATCCAGGCGAATGACTCTAAGTATGATGATTTTTTCTCAGTTCAAG 840
 Qy 841 ATAGAGAGAAATATGAATGAACAAATATTTAGAGCCCTAGGCTGAAAGATATCTTT 900

RESULT 2

AAA9294
 ID AAA9294 standard; cDNA; 1143 BP.

XX AAA9294;

XX 30-JAN-2001 (first entry)

XX Human megsin nucleotide sequence SEQ ID NO:1.

XX Megsin; mesangium-predominant gene; serpin regulated; nephropathy;
 KW IGA; immunoglobulin A; detection; renal function; renal disorder;
 KW diagnosis; biological sample; blood; urine; ss.

XX Homo sapiens.

XX WO2000571189-A1.

XX 28-SEP-2000.

XX 17-MAR-2000; 2000WO-JP01646.

XX 19-MAR-1999; 99JP-0075305.

XX 28-OCT-1999; 99JP-0306623.

XX (KURO/) KUROKAWA K.
 PA (FUSO) FUSO PHARM IND LTD.
 PA (MIYA) MIYATA T.

XX Miyata T;

XX WPI; 2000-611642/58.

XX P-PSDB; AAB24142.

XX Evaluating renal function comprises assaying megsin protein in
 PT biological sample -

XX Example 5; Page 62-66; 93pp; Japanese.

XX The present invention describes a method for evaluating renal function.
 CC The method comprises assaying megsin protein in biological sample. Also
 CC described are: (1) use of an anti-megsin protein antibody for diagnosing
 CC renal function; and (2) a kit for detecting megsin protein comprising:
 CC (a) anti-megsin protein antibody attached to solid magnetic particles;
 CC (b) direct or indirect fixing for the antibody to the particles; and
 CC (c) a magnet. The process is useful for evaluating renal function and
 CC diagnosing renal disorders by assaying megsin protein in biological
 CC samples (preferably urine or blood). The process is reproducible and
 CC gives accurate results. The present sequence encodes the human megsin

Qy 961 ATGATGCACAAATCTTACATAGAGGTCACTGAGGAGGCACCGAGGCTACTGCTGCCACA 102

Best Local Similarity	100.00; 100.00	100.00; 100.00	100.00; 100.00
Matches 1143; Conservative	0; Mismatches	0; Indels	0; Gaps

Query Match 100.0%; Score 1143; DB 22; Length 1143;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1143: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

	Db	1081 TTCCTATTGTTATCATCAGACGCATCATTCATTCATCGTGGCAAAAGTTCTTGCCCT	1141
	Dc	1141 TGA 1143	
	Dd	1141 TGA 1143	
	RESULT 4		
	ID	AQ56670 standard; cDNA; 1950 BP.	
	XX	AQ56670;	
	DT	25-MAR-2003 (updated)	
	DT	16-AUG-1994 (first entry)	
	DE	Human megakaryocyte differentiation factor coding sequence.	
	KW	Human megakaryocyte differentiation factor; MDF; thrombopoietin; haematopoietic stimulating factor; thrombocytopoemia; platelet; bone marrow transplantation; cancer chemotherapy; ds.	
	OS	Homo sapiens.	
	Key	Location/Qualifiers 74..1216 /*tag= a /product= megakaryocyte_differentiation_factor	
	PX	EP563884-Al.	
	PD	23-FEB-1994.	
	PF	19-JUL-1993; 93BP-0305654.	
	PR	17-JUL-1992; 92JP-0212305. 04-MAR-1993; 93JP-0067339.	
	PA	(SUNR) SUNFORY LTD. (TSUJ/) TSUJIMOTO Y.	
	PI	Ishida N, Iwasa F, Kurihara T, Miura K, Nakazato H; Tsujimoto M, Tsutsuka N, Yamaguchi N, Yamaichi K;	
	DR	WFI; 1994-C58782/08. P-PADB; AAR48379.	
	PT	New megakaryocyte differentiation factor - isolated from human epidermoid carcinoma cells, used to treat conditions involving a decrease in platelets	
	PS	Claim 13; Page 30-32; 47pp; English.	
	XX	The cDNA coding for human MDF was isolated from a library prepared using mRNA derived from human epidermoid carcinoma A431 cells. Human MDF stimulates differentiation of megakaryocytes from myeloid cells in the presence of IL-3. The MDF acts in vivo as a thrombopoietin making it useful for treatment of diseases involving a decrease in platelet number (esp. thrombocytopenia) such as occurs in bone marrow transplantation and in chemotherapy. (Updated on 25-MAR-2003 to correct PN field.)	
	SQ	Sequence 1950 BP; 570 A; 384 C; 407 G; 589 T; 0 other;	
	Query Match	100.0%; Score 1143; DB 15; Length 1950;	
	Best Local Similarity	100.0%; Pred. No. 0;	
	Mismatches	Conservative 0; Indels 0; Gaps 0;	
	Matches	1143; Conservative 0; Indels 0; Gaps 0;	
	QY	1 ATGCCCTCCCCTGTCTGGAGCAAATCAGAGATTTTTCTCTCACACTGTTTCAGAGATCGAT	60
	DB	74 ATGCCCTCCCCTGTCTGGAGCAAATCAGAGATTTTTCTCTCACACTGTTTCAGAGATCGAT	133

		Matches 1:43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	ATGGCTCCCTTGTGAGCAAAATCAGAGTCTTTCCTCAACCTGTTTCAGAGAGTGGAT	60		
Db	365	ATGGCTCCCTTGTGAGCAAAATCAGAGTCTTTCCTCAACCTGTTTCAGAGAGTGGAT	424		
QY	61	GACAAATCAAGAAATGAAATGTTTCTTCTCTCTGAGCTTCTGGTGGCCCTGACC	120		
Db	425	GACAAATCAAGAAATGAAATGTTTCTTCTCTCTGAGCTTCTGGTGGCCCTGACC	484		
QY	121	CTGGTCCGCTTGGCGCTCAAGATCTCCTCTCTCAGATTGATGATGATGATGAT	180		
Db	485	CTGGTCCGCTTGGCGCTCAAGATCTCCTCTCTCAGATTGATGATGATGATGAT	544		
QY	181	AACATCTCTCAGATGATGAAATCTTCTTAATAGTCAGTCAAGGCTCCAGTCTCA	240		
Db	545	AACATCTCTCAGATGATGAAATCTTCTTAATAGTCAGTCAAGGCTCCAGTCTCA	604		
QY	241	AAAGAGTCTTCTGATATAAATGATCCACAAAGATATGATCTCAGATGATGAT	300		
Db	605	AAAGAGTCTTCTGATATAAATGATCCACAAAGATATGATCTCAGATGATGAT	664		
QY	301	GGGCTTTTCTGAAAAAGTGTATGCTTTTCAATAGTCTACATGATGATGATGAT	360		
Db	665	GGGCTTTTCTGAAAAAGTGTATGCTTTTCAATAGTCTACATGATGATGATGAT	724		
QY	361	TTATAGATGCGCAAGTGGAGCGATGATTTAGATCAATTTAGAGACACTAGACCT	420		
Db	725	TTATAGATGCGCAAGTGGAGCGATGATTTAGATCAATTTAGAGACACTAGACCT	784		
QY	421	AATATTAATAGTGGTGTGAAATGAAACACATGTCGCAAAATCAAGAACTGAT	480		
Db	785	AATATTAATAGTGGTGTGAAATGAAACACATGTCGCAAAATCAAGAACTGAT	844		
QY	481	GTGGCATATGATCTCTGCTTAAGTGTCTGATGATGATGATGATGATGATGAT	540		
Db	845	GTGGCATATGATCTCTGCTTAAGTGTCTGATGATGATGATGATGATGATGAT	904		
QY	541	TGGCAATCAGCTTCCACAGAGGCAAAATCAATTTGCAATTTCAATTTCCCAATG	600		
Db	905	TGGCAATCAGCTTCCACAGAGGCAAAATCAATTTGCAATTTCAATTTCCCAATG	964		
QY	601	TCTGGGAGGCACTCGCCATGATGATCAGGAAAGGAAATGATGATGATGATGAT	660		
Db	965	TCTGGGAGGCACTCGCCATGATGATCAGGAAAGGAAATGATGATGATGATGAT	1024		
QY	661	GACCATCAATGAGATTTCTGAGTCAATGATGATGATGATGATGATGATGAT	720		
Db	1025	GACCATCAATGAGATTTCTGAGTCAATGATGATGATGATGATGATGATGAT	1084		
QY	721	CTGCTGAGATGACCTCTGAAATTTGAAACAAACTGACCTTTGAGATCTAATG	780		
Db	1085	CTGCTGAGATGACCTCTGAAATTTGAAACAAACTGACCTTTGAGATCTAATG	1144		
QY	781	TGACCAATCAAGGCAATGATGATGATGATGATGATGATGATGATGATGAT	840		
Db	1145	TGACCAATCAAGGCAATGATGATGATGATGATGATGATGATGATGATGAT	1204		
QY	841	ATAGAGAGATTTAGATGATGATGATGATGATGATGATGATGATGATGAT	900		
Db	1205	ATAGAGAGATTTAGATGATGATGATGATGATGATGATGATGATGATGAT	1264		
QY	901	GATCAATCAAGGCAATGATGATGATGATGATGATGATGATGATGATGAT	960		
Db	1265	GATCAATCAAGGCAATGATGATGATGATGATGATGATGATGATGATGAT	1324		
QY	961	ATGATCAAAATTTTATGATGATGATGATGATGATGATGATGATGATGAT	1020		
Db	1325	ATGATCAAAATTTTATGATGATGATGATGATGATGATGATGATGATGAT	1384		
QY	1021	GGAAGTAAATTTGATGATGATGATGATGATGATGATGATGATGATGAT	1080		
Db	1385	GGAAGTAAATTTGATGATGATGATGATGATGATGATGATGATGATGAT	1444		

QY	1081	TTCTATTGTTATCAGAAAGATGACATCATCTTATTCAGTGGCAAGTTTCTTGCCT	1140		
Db	1445	TTCTATTGTTATCAGAAAGATGACATCATCTTATTCAGTGGCAAGTTTCTTGCCT	1504		
QY	1141	TGA 1143			
Db	1505	TGA 1507			
RESULT 6					
ACAS560	ACA56060 standard; cDNA; 254 BP.				
ID	ACA56060				
AC	ACA56060				
DT	06-JUN-2003 (first entry)				
DE	Human signalling pathway polynucleotide probe SEQ ID NO 658.				
KW	Human; probe; ss; array element; Parkinson's disease;				
KW	signalling pathway population; cancer; adenocarcinoma; leukaemia;				
KW	immunopathy; AIDS; asthma; neuropathy; Alzheimer's disease; microarray.				
OS	Homo sapiens.				
PN	US6500938-B1.				
PD	31-DEC-2002.				
PF	30-JAN-1998; 98US-0016434.				
PR	30-JAN-1998; 98US-0016434.				
PA	(INCY-) INCYTE GENOMICS INC.				
PI	Au-Young J, Sealhamer JU;				
XX	WPI; 2003-352189/33.				
PT	Combination of polynucleotide probes, useful as array elements in a				
PT	microarray for monitoring the expression of a number of target				
PT	polynucleotides -				
PS	Claim 1; SEQ ID NO 658; 65pp; English.				
CC	The invention relates to a combination which, comprises a number of				
CC	polynucleotide probes comprising a sequence selected from one of the 1490				
CC	sequences mentioned in the specification. The combination is useful as an				
CC	array element in a microarray for monitoring the expression of a number				
CC	of target polynucleotides. The microarray is particularly useful in the				
CC	diagnosis and treatment of cancer and immunopathology and neuropathology.				
CC	The microarray is useful in diagnostics and treatment regimens, drug				
CC	discovery and development, toxicological and carcinogenicity studies,				
CC	forensics and pharmacogenomics. The microarray is also useful for				
CC	monitoring progression of diseases and for developing sophisticated				
CC	profiles for the effects of currently available therapeutic drugs. The				
CC	combination is also useful for purifying a subpopulation of mRNAs, cDNAs				
CC	and genomic fragments and in research and diagnostic applications. The				
CC	array can detect changes in expression in a large number of genes coding				
CC	for different signalling pathway populations which can be used to diagnose				
CC	various diseases including cancer e.g. adenocarcinoma and leukaemia,				
CC	immunopathies e.g. AIDS and asthma, neuropathies e.g. Alzheimer's disease				
CC	and Parkinson's disease. The present sequence represents a polynucleotide				
CC	probe of the invention.				
CC	Note: The sequence data for this patent did not form part of the printed				
CC	specification but was obtained in electronic format directly from USPTO				
CC	at seqdata.uspto.gov/sequence.html?DocID=06500938B1.				
XX	Sequence 254 BP; 83 A; 38 C; 64 G; 69 T; 0 other;				
QY	Query Watch 22.2%; Score 254; DB 25; Length 254;				
XX	Best Local Similarity 100.0%; Pred. No. 7.7e-123;				

Sequence 316 BP; 92 A; 63 C; 72 G; 88 T; 1 other;

Matches 280; Conservative 0; Mismatches 1; Indels 0; Gaps 0

822 ATTTTTCCTCAGTTCAGATAGAGAGAAATATGAAATGAAACAATATTTGAGAGCCCT 881

DB I A T T T T C C T C A G T T C A A G A T A G A G A G A A T A T G T A A T G A A C A A T A T T G A G A G C C C C T C C C

00Z AGOCCIGRRRAAFI C I I U A I U A I C O F T A S C O
QY

100

[illegible]

1003 CCGCCCTACTGCTCCACCGGAGTAATATTGTAGAAAGCAACTCCCTCAGTCCACGCT 106

240

[illegible]

QY 1062 GTTAGAGCGAGCACCACATCCATTCCTATTCGTATACAGGAGG 1102

241 GTTTAGAGCTGACCAACCCATTCCTATTGTATATCAGGAAGG 281

AAF68315/c

AAAF68315 stained; CDNA; 800 BF.
XX

AC AAF68315;
yy

DT 12-APR-2001 (first entry)

AA Human lung tumour protein related nucleotide sequence SEQ ID NO:233.
DE

XX Human: lung cancer; lung tumour; lung tumour protein; gene therapy;
KW

KW lung cancer antigen; lung tumour-specific antigen; diagnosis; vaccine;
KW carcinoembryonic antigen; antigen inhibition; ss

XX 3

[illegible]

PN
XX
WJZ001008Z8-AZ.

PD 04-JAN-2001.
yy

CC various diseases including cancer e.g. adenocarcinoma and leukaemia;
CC immunopathies e.g. AIDS and asthma, neuropathies e.g. Alzheimer's disease

CC Note: The sequence data for this patent did not form part of the printed

XX
SQ
Sequence 316 BP: 92 A: 63 C: 72 G: 88 T: 1 other:

Best Local Similarity 99.6%; Pred. No. 3.2e-110;

822 ATTTTTCCTCAGTTCAGATAGAGAGAAATATGAAATGAAACAATATTTGAGAGCCCT 881

QY 882 AGGGCTGAAAGATATCTTTGATGAATCCAAAGCAGATCTCTCTGGGATTGCTTCGGGGG 941

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

DB TGT TCGTCGTATATATCAAGGATGATGACAAATCTTACATAGAGGTCACCTGACGACGAGGCTTCTGG

Db 181 CGAGGCTACTGCTGCCACAGGAAGTAATATGTAGAAAGCAACTCCCTCAGTCCAGCT 240

D**b** 241 GTTAGAGCTGACCA^{CCCATTCCTATTTGTTATCAGGAAGG} 281

RESULT 8

XX
MC
00069375.

DI XX
12-REF-2001 (11132 CHCTY)

KW Human; lung cancer; lung tumour; lung tumour protein; gene therapy;
lung tumour antigen; diagnosis; vaccine

XX Homo sapiens. OS

XXI

PF 3C-JUN-2000; 2000WO-US18051.
XX

PR 17-DEC-1999; 99US-0466867.

PR 22-MAR-2000; 2000US-0533077.
 DE 10 APR 2000; 2000US-0546259

FR XX
03=00N=2000; 2000000-00001000.

PI Wang T, Bangur CS, Lodes MJ, Fanger GR, Vedvick TS, Carter D;

DR WPI; 2001-071488/08.
 XX Lung tumor-associated proteins and the nucleic acids that encode them,
 PT useful for preventing, diagnosing and treating lung cancer -
 XX
 XX Example 1; Page 223; 436pp; English.
 XX
 XX The present invention describes immunogenic portions of lung tumour-
 CC associated proteins (I) and the nucleic acids (NAs) that encode them.
 CC (I) have cytostatic activity and can be used in gene therapy, antisense
 CC inhibition and in vaccines. The NAs and the lung tumour-associated
 CC proteins they encode may be used in the prevention, treatment and
 CC diagnosis of diseases associated with their inappropriate expression,
 CC especially lung cancers. For example, the NAs may be administered to
 CC treat diseases by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of the protein by expressing inactive proteins
 CC or to supplement the patient's own production of (I). Additionally, the
 CC NAs may be used to produce the lung-tumour associated protein, according
 CC to standard recombinant DNA methodology. Conversely, antisense NA
 CC molecules may be administered to down regulate protein expression by
 CC binding with the cells own genes and preventing their expression. The NA
 CC and complementary sequences may also be used as DNA probes in diagnostic
 CC assays to detect and quantitate the presence of similar NA sequences in
 CC samples, and hence which patients may be in need of treatment for lung
 CC cancer. The (I) may be used as antigens in the production of antibodies
 CC and in assays to identify modulators (agonists and antagonists) of the
 CC expression and activity of the protein. AAF68083 to AAF68878 and
 CC AAB76848 to AAB76878 represent human lung tumour protein related
 CC nucleotide and protein sequences which are used in the exemplification
 CC of the present invention.
 XX
 XX Sequence 600 BP; 198 A; 106 C; 135 G; 161 T; 0 other;
 SQ
 Query Match 10.6%; Score 121; DB 22; Length 600;
 Best Local Similarity 100.0%; Pred. No. 6.6e-53;
 Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1023 AAGTAATATTGTAGAAAGCAACTCCCTCAGTCACCGCTGTTAGAGCTGACCAACCAT 1082
 Db 600 AAGTAATATTGTAGAAAGCAACTCCCTCAGTCACCGCTGTTAGAGCTGACCAACCAT 541
 QY 1083 CCTATTGTTATCAGGAGGATGACATCTATTTCAGTGGCAAGTTCTTGCCCTTG 1142
 Db 540 CCTATTGTTATCAGGAGGATGACATCTATTTCAGTGGCAAGTTCTTGCCCTTG 491
 QY 1143 A 1143
 Db 480 A 480
 RESULT 9
 ABK38226/c
 ID ABK38226 standard; cDNA; 600 BP.
 XX
 XX AC ABK38226;
 XX
 XX 21-MAY-2002 (first entry)
 XX
 XX cDNA encoding clone #22045 of lung tumour protein.
 DE
 XX Lung tumour; cancer; T cell; immune response stimulator;
 KW cytostatic; gene; ss.
 XX
 XX Homo sapiens.
 OS
 XX WC200204514-A2.
 PN
 XX 17-JAN-2002.
 PD
 XX 10-JUL-2001; 2001WO-US22058.
 XX
 XX 11-JUL-2000; 2000US-0614124.
 PR
 XX 29-AUG-2000; 2000US-0651563.
 PF

ER 08-SEP-2000; 2000US-0658824.
 PR 26-SEP-2000; 2000US-0671325.
 PR 06-OCT-2000; 2000US-0677419.
 PR 30-OCT-2000; 2000US-0702705.
 PR 13-DEC-2000; 2000US-0736457.
 PR 03-MAY-2001; 2001US-0849626.
 XX
 XX (CORI-) CORIYA CORP.
 PA
 XX Wang T, Watanabe Y, Henderson RA, Johnson JC, Retter MW;
 XX Marnerakis M, Carter D, Fanger GR, Vedvick TS, Bangur CS;
 PI Mcnabb A, Wang A, Fanger N, Switzer A, Mcneill PD, Clapper JD;
 XX
 XX WPI; 2002-164634/21.
 XX
 XX Novel polynucleotide encoding a lung tumour polypeptide useful for
 PT stimulating and/or expanding T cells specific for a tumour protein -
 PT
 XX Example 1; SEQ ID No 233; 223pp; English.
 PS
 XX The invention describes an isolated polynucleotide and polypeptide
 CC useful for stimulating and/or expanding T cells specific for a tumour
 CC protein for determining the presence of a cancer in a patient. A
 CC composition containing the polynucleotide and/or polypeptide is useful
 CC for treating a lung cancer in a patient. The polypeptide is useful for
 CC removing tumour cells from a biological sample. The polynucleotide is
 CC also useful as probe or primer to detect the level of mRNA encoding a
 CC tumour protein. This sequence encodes a lung tumour associated protein
 CC or protein fragment, described in the method of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 600 BP; 198 A; 106 C; 135 G; 161 T; 0 other;
 SQ
 Query Match 10.6%; Score 121; DB 24; Length 600;
 Best Local Similarity 100.0%; Pred. No. 6.6e-53;
 Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1023 AAGTAATATTGTAGAAAGCAACTCCCTCAGTCACCGCTGTTAGAGCTGACCAACCAT 1082
 Db 600 AAGTAATATTGTAGAAAGCAACTCCCTCAGTCACCGCTGTTAGAGCTGACCAACCAT 541
 QY 1083 CCTATTGTTATCAGGAGGATGACATCTATTTCAGTGGCAAGTTCTTGCCCTTG 1142
 Db 540 CCTATTGTTATCAGGAGGATGACATCTATTTCAGTGGCAAGTTCTTGCCCTTG 481
 QY 1143 A 1143
 Db 480 A 480
 RESULT 10
 ACA10555/c
 ID ACA10555 standard; cDNA; 600 BP.
 XX
 XX AC ACA10555;
 XX
 XX 05-JUN-2003 (first entry)
 XX
 XX Human lung cancer-associated cDNA, SEQ ID 233.
 DE
 XX Human; lung cancer; ss; lung tumour; cytostatic; vaccine;
 KW T cell expansion; CD4; CD8.
 XX
 XX Homo sapiens.
 OS
 XX US2002197669-A1.
 PN
 XX 26-DEC-2002.
 PD
 XX 03-MAY-2001; 2001US-0849626.
 PF

13-DEC-2000; 2000US-0736457.

(BANG/) BANGUR C S.
(FANG/) FANGER G R.
(WANG/) WANG A.
(SWIT/) SWITZER A P.
(MCNE/) MCNEILL P D.
(CLAP/) CLAPPER J D.

Bangur CS, Fanger GR, Wang A, Wang T, Switzer AP, McNeill PD;
Clapper JD;
WPI; 2003-352750/33.

Novel lung cancer polynucleotide encoding lung cancer protein, useful
for detecting the presence of lung cancer in a patient, and in
pharmaceutical compositions, e.g. vaccines, for treating lung cancer -
Example 1; Page -; 72pp; English.

The invention relates to a polynucleotide encoding a lung tumour protein,
comprising a sequence selected from any of the 14 sequences
mentioned in the specification, or a sequence (S2) mentioned in
specification, complement of S1, sequences consisting of at least 20
contiguous residues of S1, sequences that hybridise to S1, sequences
having 75%, preferably 90%, identity to S1, or degenerate variants of
S1. Also included are an isolated polypeptide (comprising a sequence (S3)
selected from any one of the 4 amino acid sequences mentioned in the
specification, a sequence encoded by the polynucleotide, or sequences
having at least 70%, preferably 90%, identity to a sequence encoded by
the polynucleotide), an expression vector comprising the polynucleotide
operably linked to an expression control sequence, a host cell
transformed or transfected with the vector, an isolated antibody (or its
antigen-binding fragment) that specifically binds to the polypeptide,
detecting the presence of a cancer in a patient, a fusion protein
comprising the polypeptide, an oligonucleotide that hybridises to
S1 under moderately stringent conditions, stimulating and/or expanding T
cells specific for a tumour protein (comprising contacting T cells with
the polynucleotide, protein or antigen-presenting cells, under conditions
and for a time sufficient to permit the stimulation and/or expansion of T
cells; and inhibiting the development of a cancer in a patient (by
incubating CD4+ and/or CD8+ T cells isolated from a patient with the
polynucleotide, protein or antigen presenting cells that express the
polynucleotide, such that T cells proliferate, administering to the
patient an effective amount of the proliferated T cells, and thus
inhibiting the development of a cancer in the patient. The
polynucleotide, protein and cells are useful in a composition for
stimulating an immune response in a patient, and for treating a cancer in
a patient (particularly lung cancer). The oligonucleotide is useful for
determining the presence of a cancer in a patient. The protein and
oligonucleotides are useful in pharmaceutical compositions, e.g.
vaccines. The polynucleotide is also useful as a probe or primer for
nucleic acid hybridisation, and in the design and preparation of
ribozyme molecules for inhibiting expression of tumour polypeptides and
proteins in tumour cells. An amplified portion of the polynucleotide is
useful for isolating a full-length gene from a suitable library.
The present sequence is a cDNA (full length, extended or partial)
isolated from a library derived from lung tumour/cancer cells.
Note: The sequence data for this patent did not form part
of the printed specification, but was obtained in electronic
format directly from the USPTO
at seqdata.uspto.gov/sequence.html?DocId=20020137669.

Sequence 600 BP; 198 A; 106 C; 135 G; 161 T; 0 other;
Query Match 10.6%; Score 121; DB 25; Length 600;
Best Local Similarity 100.0%; Pred. No. 6.6e-53;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1023 AAGTAAATTGTAAGAAAGCACTCCCTCAGTCCAGCTGTTAGAGTGACCAACCCATT 1082
|||||

Db 600 AAGTAAATTGTAAGAAAGCACTCCCTCAGTCCAGCTGTTAGAGTGACCAACCCATT 541
Qy 1083 CCTATTGTTATCAGGAAGGATGACATCATCTATTCTGCGAAAGTTCTTSCCCTTG 1142
Db 540 CCTATTGTTATCAGGAAGGATGACATCATCTATTCTGCGAAAGTTCTTSCCCTTG 461
Qy 1143 A 1143
Db 480 A 480

RESULT 11
AEX99506/c
ID AEX99506 standard; cDNA; 600 BP.
XX AC AEX99506;
XX DT 22-MAY-2003 (first entry)
XX DE Lung cancer therapy and diagnosis associated cDNA #233.
XX KW Lung cancer; cytostatic; vaccine; gene therapy; cancer;
XX gene; ss.
XX OS Homo sapiens.
XX PN US2002172952-A1.
XX PD 21-NOV-2002.
XX PF 10-JUL-2001; 2001US-0902941.
XX PR 30-JUN-1999; 99US-0346492.
XX PR 15-OCT-1999; 99US-0419356.
XX PR 17-DEC-1999; 99US-0466867.
XX PR 30-DEC-1999; 99US-0476300.
XX PR 06-MAR-2000; 2000US-0519642.
XX PR 22-MAR-2000; 2000US-0533077.
XX PR 10-APR-2000; 2000US-0546259.
XX PR 27-APR-2000; 2000US-0589184.
XX PR 05-JUN-2000; 2000US-0614124.
XX PR 11-JUL-2000; 2000US-0651563.
XX PR 08-SEP-2000; 2000US-0658824.
XX PR 26-SEP-2000; 2000US-0671325.
XX PR 06-OCT-2000; 2000US-0677419.
XX PR 30-OCT-2000; 2000US-0702705.
XX PR 13-DEC-2000; 2000US-0736457.
XX PR 03-MAY-2001; 2001US-0849626.
XX (CORI-) CORIXA CORP.
XX Henderson RA, Wang T, Watanabe Y, Johnson JC, Retter MW, Durham M;
PI Carter D, Fanger GR, Vedvick TS, Bangur CS, McNabb A;
XX WPI; 2003-328427/31.
XX New polynucleotide, useful for preparing a composition for treating or
PT inhibiting development of cancer, e.g. lung cancer -
XX Example 1; SEQ ID NO 233; 82pp; English.
XX The invention describes an isolated polynucleotide comprising one of 32
CC sequences, complement or degenerate variants of them. The polynucleotide
CC is useful for preparing a composition e.g. a vaccine or for gene therapy,
CC for treating or inhibiting development of cancer, e.g. lung cancer.
CC This sequence represents a polynucleotide associated with the
CC compositions and methods for the therapy and diagnosis of lung cancer.
XX Sequence 600 BP; 198 A; 106 C; 135 G; 161 T; 0 other;
SQ Query Match 10.6%; Score 121; DB 25; Length 600;
Best Local Similarity 100.0%; Pred. No. 6.6e-53;

Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1023 AACTAATTTGTAGAAAGCAACTCCCTCAGTCCACGCTGTTTAGAGTGCACCAACCAT 1082
 DB 600 AAGTAATATTGTAGAAAGCAACTCCCTCAGTCCACGCTGTTTAGAGTGCACCAACCAT 541

QY 1083 CCATATTGTTATCAGGAAGGATGACATCATCTTTTCAGTGGCAAAAGTTTTCGCCCTTG 1142
 DB 540 CCATATTGTTATCAGGAAGGATGACATCATCTTTTCAGTGGCAAAAGTTTTCGCCCTTG 481

QY 1143 A 1143
 DB 480 A 480

RESULT 12
 AAX56727/c
 ID AAX56727 standard; DNA; 42 BP.
 XX AC AAX56727;
 DT 14-JUL-1999 (first entry)
 XX WO99-5652 primer 13.
 XX Megsin; mesangial cell; treatment; diagnosis; disease; IGA nephropathy;
 XX human; rat; murine; primer; ss.
 XX Synthetic.
 OS WO9915652-A1.
 PN 01-APR-1999.
 XX 22-SEP-1998; 98WO-JP04269.
 XX 22-SEP-1997; 97JP-0275302.
 XX (KURO/) KUROKAWA K.
 XX (MIYA/) MIYATA T.
 XX Miyata T;
 PI WPI; 1999-276983/23.
 DR P-PSDB; AAY08256.
 XX Megsin protein expressed specifically in mesangial cells
 PS Claim 4; Page 72-76; 100pp; Japanese.
 CC This invention describes the isolation of novel megsin nucleic acid and
 CC proteins from human, rat and mouse tissue. This protein is expressed
 CC specifically in mesangial cells. The products of the invention are
 CC useful for the treatment and diagnosis of diseases involving mesangial
 CC cells, such as IGA nephropathy. AAX56715-X56752 represent primers used
 CC in the method of the invention.

QY 1113 CTTATTCACTGCGCAAGTTTCTTGCCCTTGA 1143
 DB 42 CTTATTCACTGCGCAAGTTTCTTGCCCTTGA 12

RESULT 13
 AAX56714
 ID AAX56714 standard; DNA; 1147 BP.
 XX AC AAX56714;
 XX

Query Match 2.7%; Score 31; DB 20; Length 42;
 Best Local Similarity 100.0%; Pred. No. 1.1e-05;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1113 CTTATTCACTGCGCAAGTTTCTTGCCCTTGA 1143
 DB 42 CTTATTCACTGCGCAAGTTTCTTGCCCTTGA 12

RESULT 14
 AAX55239
 ID AAC55239 standard; DNA; 1147 BP.
 XX AC AAC55239;
 XX 30-JAN-2001 (first entry)
 XX Mouse megsin nucleotide sequence SEQ ID NO:20.
 XX Megsin; mesangium-predominant gene; serpin regulated; nephropathy;
 XX IGA; immunoglobulin A; detection; renal function; renal disorder;
 XX diagnosis; biological sample; blood; urine; ss.
 XX Mus musculus.
 XX WC2000057189-A1.
 XX 28-SEP-2000.
 XX 17-MAR-2000; 2000WC-JP01646.

DT 14-JUL-1999 (first entry)
 XX Mouse megsin DNA.
 XX Megsin; mesangial cell; treatment; diagnosis; disease; IGA nephropathy;
 XX human; rat; murine; ss.
 XX Mus musculus.
 OS Key
 FT CDS
 FT Location/Qualifiers
 FT 1..1107
 FT /*tag= a
 FT /product= "megsin"
 FT /note= "partial sequence, no start codon given"
 XX WO9915652-A1.
 XX 01-APR-1999.
 XX 22-SEP-1998; 98WO-JP04269.
 XX 22-SEP-1997; 97JP-0275302.
 XX (KURO/) KUROKAWA K.
 XX (MIYA/) MIYATA T.
 XX Miyata T;
 PI WPI; 1999-276983/23.
 DR P-PSDB; AAY08256.
 XX Megsin protein expressed specifically in mesangial cells
 PS Claim 4; Page 72-76; 100pp; Japanese.
 CC This invention describes the isolation of novel megsin nucleic acid and
 CC proteins from human, rat and mouse tissue. This protein is expressed
 CC specifically in mesangial cells. The products of the invention are
 CC useful for the treatment and diagnosis of diseases involving mesangial
 CC cells, such as IGA nephropathy.
 XX Sequence 1147 BP; 351 A; 239 C; 250 G; 307 T; 0 other;
 SQ

Query Match 2.7%; Score 31; DB 20; Length 1147;
 Best Local Similarity 100.0%; Pred. No. 1.4e-05;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 832 CAGTTCAAGATAGAGAGAAATTATGAAATGA 862
 DB 796 CAGTTCAAGATAGAGAGAAATTATGAAATGA 826

RESULT 14
 AAC55239
 ID AAC55239 standard; DNA; 1147 BP.
 XX AC AAC55239;
 XX 30-JAN-2001 (first entry)
 XX Mouse megsin nucleotide sequence SEQ ID NO:20.
 XX Megsin; mesangium-predominant gene; serpin regulated; nephropathy;
 XX IGA; immunoglobulin A; detection; renal function; renal disorder;
 XX diagnosis; biological sample; blood; urine; ss.
 XX Mus musculus.
 XX WC2000057189-A1.
 XX 28-SEP-2000.
 XX 17-MAR-2000; 2000WC-JP01646.

XX 19-MAR-1999; 99JP-0075305.
PR 28-OCT-1999; 99JP-0306623.
XX (KURO/) KUROKAWA K.
PA (FUSO) FUSO PHARM IND LTD.
PA (MIYA/) MIYATA T.
XX Miyata T;
PI
XX WPI; 2000-511642/58.
DR P-PSDB; AA324151.
XX
PT Evaluating renal function comprises assaying megin protein in
PT biological sample -
XX
PS Disclosure; Page 85-88; 93pp; Japanese.
XX
CC The present invention describes a method for evaluating renal function.
CC The method comprises assaying megin protein in biological sample. Also
CC described are: (1) use of an antimegin protein antibody for diagnosing
CC renal function; and (2) a kit for detecting megin protein comprising:
CC (a) antimegin protein antibody attached to solid magnetic particles;
CC (b) direct or indirect fixing for the antibody to the particles; and
CC (c) a magnet. The process is useful for evaluating renal function and
CC diagnosing renal disorders by assaying megin protein in biological
CC samples (preferably urine or blood). The process is reproducible and
CC gives accurate results. The present sequence encodes the mouse megin
CC protein, which is given in the exemplification of the present invention.
XX
SQ Sequence 1147 BP; 351 A; 239 C; 250 G; 307 T; 0 other;

Query Match 2.7%; Score 31; DB 21; Length 1147;
Best Local Similarity 100.0%; Pred. No. 1.4e-05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 832 CAGTTCAGATAGAGAGAGATTATGAAATGA 862
Db 796 CAGTTCAGATAGAGAGAGATTATGAAATGA 826

RESULT 15
AAF82440
ID AAF82440 standard; cDNA; 1147 BP.
XX
AC AAF82440;
XX
DT 10-JUL-2001 (first entry)
XX
DE Mouse megin nucleotide sequence.
XX
KW Mouse; megin; mesangial cell proliferative nephritis; nephrotropic;
KW transgenic mouse; glomerular disease; animal model; drug screening; ss.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT CDS 1..1107
FT */*tag= a*
FT */product= "megsin"*
XX
FN WO200124628-A1.
XX
PD 12-APR-2001.
XX
PP 06-OCT-2000; 2000WO-JP06988.
XX
PR 06-OCT-1999; 99JP-0285736.
XX (KURO/) KUROKAWA K.
PA (MIYA/) MIYATA T.
XX
PI Miyata T;

XX WPI; 2001-300136/31.
DR P-PSDB; AA83077.
XX
PT Mouse model for mesangial cell proliferative nephritis for development
PT and screening of new treatments -
XX
XX Disclosure; Page 50-52; 62pp; Japanese.
XX
CC The present sequence encodes mouse megin. The human megin coding
CC sequence may be introduced into a mouse to produce an animal model of
CC mesangial cell proliferative nephritis. The symptoms include
CC enlargement of the mesangial base region, sedimentation of an immune
CC complex and an increase in mesangial cells. The animal model is useful
CC for analysing the pathology of chronic glomerular diseases and for
CC screening compositions for prevention and treatment of the diseases.
CC Highly uniform models can be made easily and in large numbers using
CC this method.
XX
SQ Sequence 1147 BP; 351 A; 239 C; 250 G; 307 T; 0 other;

Query Match 2.7%; Score 31; DB 22; Length 1147;
Best Local Similarity 100.0%; Pred. No. 1.4e-05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 832 CAGTTCAGATAGAGAGAGATTATGAAATGA 862
Db 796 CAGTTCAGATAGAGAGAGATTATGAAATGA 826

Search completed: December 14, 2003, 04:48:33
Job time : 381 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2003, 02:33:04 ; Search time 426 Seconds
(without alignments)
8917.544 Million cell updates/sec

Title: US-09-508-997A-1
Perfect score: 1143
Sequence: 1 atggcctccctgtgcgcg.....Gcaagtttttgccttga 1143

Scoring table: OLIGO NUC
Gapop 50.0 , Gapext 60.0

Searched: 2201672 seqs, 1661799599 residues

Word size: 0
Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

- Database : Published Applications NA:*
- 1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
 - 2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq:*
 - 3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq:*
 - 4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
 - 5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:*
 - 6: /cgn2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq:*
 - 7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:*
 - 8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:*
 - 9: /cgn2_6/prodata/2/pubpna/US09A_PUBCOMB.seq:*
 - 10: /cgn2_6/prodata/2/pubpna/US09B_PUBCOMB.seq:*
 - 11: /cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq:*
 - 12: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq:*
 - 13: /cgn2_6/prodata/2/pubpna/US09_PUBCOMB.seq:*
 - 14: /cgn2_6/prodata/2/pubpna/US10A_PUBCOMB.seq:*
 - 15: /cgn2_6/prodata/2/pubpna/US10B_PUBCOMB.seq:*
 - 16: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq:*
 - 17: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:*
 - 18: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1143	100.0	1950	9	US-09-140-719-30
2	1143	100.0	1950	14	US-10-091-442-30
3	399	34.9	399	13	US-10-029-386-21145
4	399	34.9	592	13	US-10-029-386-7445
5	121	10.6	521	13	US-10-029-386-10599
6	121	10.6	600	10	US-09-736-457-233
7	121	10.6	600	10	US-09-902-941-233
8	121	10.6	600	10	US-09-849-626-233
9	121	10.6	600	11	US-09-476-300-233
10	121	10.6	600	13	US-10-113-872-233
11	121	10.6	600	15	US-10-017-754-233
12	119	10.4	119	13	US-10-029-386-24299
13	25	2.2	37	9	US-09-140-719-32
14	25	2.2	37	14	US-10-091-442-32
15	23	2.0	23	9	US-09-140-719-14

16	23	2.0	23	14	US-10-091-442-14	Sequence 14, Appl
17	21	1.8	21	9	US-09-140-719-15	Sequence 15, Appl
18	21	1.8	21	9	US-09-140-719-16	Sequence 16, Appl
19	21	1.8	21	9	US-09-140-719-17	Sequence 17, Appl
20	21	1.8	21	9	US-09-140-719-18	Sequence 18, Appl
21	21	1.8	21	9	US-09-140-719-19	Sequence 19, Appl
22	21	1.8	21	14	US-10-091-442-15	Sequence 15, Appl
23	21	1.8	21	14	US-10-091-442-16	Sequence 16, Appl
24	21	1.8	21	14	US-10-091-442-17	Sequence 17, Appl
25	21	1.8	21	14	US-10-091-442-18	Sequence 18, Appl
26	21	1.8	21	14	US-10-091-442-19	Sequence 19, Appl
27	21	1.8	21	14	US-10-091-442-20	Sequence 20, Appl
28	21	1.8	21	14	US-10-091-442-21	Sequence 21, Appl
29	21	1.8	21	14	US-10-091-442-22	Sequence 22, Appl
30	20	1.7	343	10	US-09-878-574-2608	Sequence 26, Appl
31	20	1.7	545	13	US-10-027-632-202376	Sequence 202376, Ap
32	20	1.7	545	14	US-10-027-632-202376	Sequence 202376, Ap
33	20	1.7	83450	11	US-09-811-469-3	Sequence 3, Appli
34	20	1.7	83450	13	US-10-370-659-3	Sequence 3, Appli
35	19	1.7	211	9	US-09-864-761-18550	Sequence 18550, A
36	19	1.7	351	10	US-09-783-590-7053	Sequence 7053, Ap
37	19	1.7	448	13	US-09-814-353-14497	Sequence 14497, A
38	19	1.7	470	9	US-09-864-761-1798	Sequence 1798, Ap
39	19	1.7	478	15	US-10-239-423-5	Sequence 5, Appli
40	19	1.7	502	13	US-09-814-353-1763	Sequence 1763, Ap
41	19	1.7	502	13	US-09-814-353-8113	Sequence 8113, Ap
42	19	1.7	613	15	US-10-239-423-7	Sequence 7, Appli
43	19	1.7	686	13	US-10-027-632-134774	Sequence 134774, A
44	19	1.7	686	14	US-10-027-632-134774	Sequence 134774, A
45	19	1.7	1293	13	US-10-027-632-124492	Sequence 124492, A

ALIGNMENTS

RESULT 1
US-09-140-719-30
; Sequence 30, Application US/09140719
; Patent No. US20010026931A1
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Fuyuki
; APPLICANT: TSUROOKA, No. US20010026931A1uc
; APPLICANT: NAKAZATO, Hiroshi
; APPLICANT: MIURA, Kenju
; APPLICANT: ISHIDA, No. US20010026931A1uhiro
; APPLICANT: KURIHARA, Tatsuya
; APPLICANT: YAMAUCHI, Kozo
; APPLICANT: YAMAGUCHI, No. US20010026931A1omi
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/140,719
; FILING DATE: 08-AUG-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/474,661
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993

QY	541	TGGCAATCAGCCTTCAACAGAGCGGAACCATAAATTCGCATTTCAAAATCTCCCAAGTGC	600
DB	614	TGGCAATCAGCCTTCAACAGAGCGGAACCATAAATTCGCATTTCAAAATCTCCCAAGTGC	673
QY	601	TCTGGGAAGCGAGTCGCCATGATGCAATCAGGAACGGAAGTTCAAATTTGCTGTATTAGAG	660
DB	574	TCTGGGAAGCGAGTCGCCATGATGCAATCAGGAACGGAAGTTCAAATTTGCTGTATTAGAG	733
QY	661	GAACCATCAATGAAGATCTTGAAGTTCAGATACAAATGCTGGCATAAACATGATGAGTTCG	720
DB	734	GAACCATCAATGAAGATCTTGAAGTTCAGATACAAATGCTGGCATAAACATGATGAGTTCG	793
QY	721	CTGCGCTGAGAATGACCTCTCTGAAATTGAAAAAAGAACTGACCTTTCAGAAATCTAATGAA	780
DB	794	CTGCGCTGAGAATGACCTCTCTGAAATTGAAAAAAGAACTGACCTTTCAGAAATCTAATGAA	853
QY	781	TGGACCAATCAAGCGGAATGACCTCTGAAGTATCTTGAAGTATCTTCTCAGTTCGAAG	840
DB	854	TGGACCAATCAAGCGGAATGACCTCTGAAGTATCTTGAAGTATCTTCTCAGTTCGAAG	913
QY	841	ATAGAGAGGAATATGAAATGAAACAAATATTTGAGAGCCCTAGGGCTGAAAGATATCTTT	900
DB	914	ATAGAGAGGAATATGAAATGAAACAAATATTTGAGAGCCCTAGGGCTGAAAGATATCTTT	973
QY	901	GATGAATCCAAAGCAGATCTCTCTGGGAATGCTTCGGGGGGTCTGTGTATATATCAGG	960
DB	974	GATGAATCCAAAGCAGATCTCTCTGGGAATGCTTCGGGGGGTCTGTGTATATATCAGG	1033
QY	961	ATGATGCAAAATCTTACATAGAGGTCATGAGAGGAGGCGACCGAGGCTACTGCTGCACA	1020
DB	1034	ATGATGCAAAATCTTACATAGAGGTCATGAGAGGAGGCGACCGAGGCTACTGCTGCACA	1093
QY	1021	GGAAGTAATATTGTGAAAGAACAATCCCTCAGTCCAGCTGTTTACAGTGTGACCAACCA	1080
DB	1094	GGAAGTAATATTGTGAAAGAACAATCCCTCAGTCCAGCTGTTTACAGTGTGACCAACCA	1153
QY	1081	TTCTATTTGTTATCAGGAGGAGTACATCATCTTATTCAGTGGCAAGTTTCTTGCCCT	1140
DB	1154	TTCTATTTGTTATCAGGAGGAGTACATCATCTTATTCAGTGGCAAGTTTCTTGCCCT	1213
QY	1141	TGA	1143
DB	1214	TGA	1216
RESULT 2			
US-10-091-442-30			
: Sequence 30, Application US/10091442			
: Publication No. US20020164711A1			
GENERAL INFORMATION:			
: APPLICANT: TSUJIMOTO, Masafumi			
: IWASA, Fuyuki			
: TSURUOKA, No. US20020164711A1uo			
: NAKAZATO, Hitoshi			
: MIURA, Kenju			
: ISHIDA, No. US20020164711A1ubiro			
: KURIHARA, Tatsuya			
: YAMAICHI, Kozo			
: YAMAGUCHI, No. US20020164711A1omi			
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR			
: NUMBER OF SEQUENCES: 34			
: CORRESPONDENCE ADDRESS:			
: ADDRESSEE: Burns, Doane, Swecker & Mathis			
: STREET: P.O. Box 1404			
: CITY: Alexandria			
: STATE: Virginia			
: COUNTRY: United States			
: ZIP: 22313-1404			
COMPUTER READABLE FORM:			
: MEDIUM TYPE: Floppy disk			
: COMPUTER: IBM PC compatible			
: OPERATING SYSTEM: PC-DOS/MS-DOS			
: SOFTWARE: PatentIn Release #1.0, Version #1.30			

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/091,442
 FILING DATE: 07-Mar-2002
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/140,719
 FILING DATE: 08-AUG-1998
 APPLICATION NUMBER: US 08/474,661
 FILING DATE: 07-JUN-1995
 APPLICATION NUMBER: US 08/091,028
 FILING DATE: 14-JUL-1993
 APPLICATION NUMBER: JP 4-212305
 FILING DATE: 17-JUL-1992
 APPLICATION NUMBER: JP 6-067339
 FILING DATE: 04-MAR-1993

ATTORNEY/AGENT INFORMATION:

NAME: McGowan, Malcolm K.
 REGISTRATION NUMBER: 39,300
 REFERENCE/DOCKET NUMBER: 001560-247
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 836-6620
 TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 30:

SEQUENCE CHARACTERISTICS:
 LENGTH: 1950 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: Double
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 STRAIN: A431

FEATURE:

NAME/KEY: misc_feature
 LOCATION: 1
 OTHER INFORMATION: /note= "DNA coding for human megakaryocyte differentiation factor."

FEATURE:

NAME/KEY: CDS
 LOCATION: 74..1217
 SEQUENCE DESCRIPTION: SEQ ID NO: 30:

US-10-091-442-30

Query Match 100.0%; Score 1143; DB 14; Length 1950;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGGCTCCCTTGTGAGCAAAATGAGAGTTTGGTTCAACCTGTTTCAGAGAGATGGAT 60
 74 ATGGCTCCCTTGTGAGCAAAATGAGAGTTTGGTTCAACCTGTTTCAGAGAGATGGAT 133
 61 GACATCAAGAAATGAAATGTGTTTCTCTCTGAGGCTCTTGGCTGCTGCTGCTGCTG 120
 134 GACAATCAAGAAATGAAATGTGTTTCTCTCTGAGGCTCTTGGCTGCTGCTGCTGCTG 193
 121 CTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
 194 CTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 253
 181 AACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
 254 AACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 313
 241 AAAAGAGTTTTTCTGATATAAATGATCCCAAGATATGATCTCAGCATTTGTAAT 300
 314 AAAAGAGTTTTTCTGATATAAATGATCCCAAGATATGATCTCAGCATTTGTAAT 373
 301 GGCCTTTTGTGAAAGTGATGCTTCTAAGAGTACATTTAGAGTGTGCGGAA 360
 374 GGCCTTTTGTGAAAGTGATGCTTCTAAGAGTACATTTAGAGTGTGCGGAA 433
 361 TTATACGATGCCAAGTGGAGGAGTTGACTTTACGAATCAATTTAGAGACACTAGAGT 420

434 TTATACGATGCCAAGTGGAGGAGTTGACTTTACGAATCAATTTAGAGACACTAGAGT 493
 421 AATATTATTAAGTGGTTGAAAATGAAAACACATGCGAATAATCAGACGTTGTTGGTCAA 480
 494 AATATTATTAAGTGGTTGAAAATGAAAACACATGCGAATAATCAGACGTTGTTGGTCAA 553
 481 GGTGACATAGCTCATCTGCTGTAATGCTGCTGTAATGCTGTAATGCTGTAATGCTGTA 540
 554 GGTGACATAGCTCATCTGCTGTAATGCTGTAATGCTGTAATGCTGTAATGCTGTAATG 613
 541 TGGCAATCAGCTTCCACCAAGAGGAAACATATAATGCTGTAATGCTGTAATGCTGTAATG 600
 614 TGGCAATCAGCTTCCACCAAGAGGAAACATATAATGCTGTAATGCTGTAATGCTGTAATG 673
 601 TGTGGAGGAGCTGCGCATGATGCTGTAATGCTGTAATGCTGTAATGCTGTAATGCTGTA 660
 674 TGTGGAGGAGCTGCGCATGATGCTGTAATGCTGTAATGCTGTAATGCTGTAATGCTGTA 733
 661 GACCATCAATGAAGATTCTTGAGCTCAGATACATATGCTGTAATGCTGTAATGCTGTAATG 720
 734 GACCATCAATGAAGATTCTTGAGCTCAGATACATATGCTGTAATGCTGTAATGCTGTAATG 793
 721 CTGCTGAGAAATGACCTCTCTCAAAATGAAAACAACTGACCTTTCAGATCTAATGAA 780
 794 CTGCTGAGAAATGACCTCTCTCAAAATGAAAACAACTGACCTTTCAGATCTAATGAA 853
 781 TGGCAATCAGCTTCCACCAAGAGGAAACATATAATGCTGTAATGCTGTAATGCTGTAATG 840
 854 TGGCAATCAGCTTCCACCAAGAGGAAACATATAATGCTGTAATGCTGTAATGCTGTAATG 913
 841 ATAGAGAGAAATTAATGAAATGAAAACATATAATGCTGTAATGCTGTAATGCTGTAATG 900
 914 ATAGAGAGAAATTAATGAAATGAAAACATATAATGCTGTAATGCTGTAATGCTGTAATG 973
 901 GATGATCAGAAAGCAGATCTCTCTGGATTGCTTCCGGGGGCTGCTGTAATGTAATCAAG 960
 974 GATGATCAGAAAGCAGATCTCTCTGGATTGCTTCCGGGGGCTGCTGTAATGTAATCAAG 1033
 961 ATGATGCAAAATCTTACATAGAGGTCATGAGGAGGCGACCGAGGCTTACTGCTGCCACA 1020
 1034 ATGATGCAAAATCTTACATAGAGGTCATGAGGAGGCGACCGAGGCTTACTGCTGCCACA 1093
 1021 GGAAGTAATATTGTAGAAAAGCAACTCCCTCAGTCACACGCTGTTTAGAGCTGACCAACCA 1080
 1094 GGAAGTAATATTGTAGAAAAGCAACTCCCTCAGTCACACGCTGTTTAGAGCTGACCAACCA 1153
 1081 TTCTATTGTTTATCAGGAGGATGACATCATCTTATTGAGTGGCAAGTTTCTTGGCTT 1140
 1154 TTCTATTGTTTATCAGGAGGATGACATCATCTTATTGAGTGGCAAGTTTCTTGGCTT 1213
 1141 TGA 1143
 1214 TGA 1216

RESULT 3

US-10-029-386-21145
 ; Sequence 21145, Application US/10029386
 ; Publication No. US20030194704A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharron G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL F
 ; FILE REFERENCE: AEMICA-X-2
 ; CURRENT FILING DATE: 2001-12-20
 ; CURRENT APPLICATION NUMBER: US/10/029,386
 ; NUMBER OF SEQ ID NOS: 34288
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 21145
 ; LENGTH: 399
 ; TYPE: DNA

ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO CHR18.3
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2
OTHER INFORMATION: SWISSPROT HIT: 002739, EVALUATION 5.00e-24
OTHER INFORMATION: NT HIT: D88575.1, EVALUATION 0.00e+00
US-10-029-386-7445

Query Match 34.9%; Score 399; DB 13; Length 592;
Best Local Similarity 100.0%; Pred. No. 3e-203;
Matches 399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 745 ATTGAAACAAACTGACCTTTTCAGATCTTAATGATGAGCAATCCAGGCGAATGACC 804
DB 146 ATTGAAACAAACTGACCTTTTCAGATCTTAATGATGAGCAATCCAGGCGAATGACC 205
QY 805 TCTAAGTATGTTGAGGTATTTTTCCTCAGTTCAGATGAGCAATCCAGGCGAATGACC 265
DB 206 TCTAAGTATGTTGAGGTATTTTTCCTCAGTTCAGATGAGCAATCCAGGCGAATGACC 265
QY 865 CAATATTGAGAGCCCTAGGGCTGAAAGATATCTTTGATGATCCAAAGCAGATCTCT 924
DB 266 CAATATTGAGAGCCCTAGGGCTGAAAGATATCTTTGATGATCCAAAGCAGATCTCT 325
QY 925 GGGATTGCTCGGGGGTCTCTGTATATATCAAGATGATGAGCAATCCAGGCGAATGACC 984
DB 326 GGGATTGCTCGGGGGTCTCTGTATATATCAAGATGATGAGCAATCCAGGCGAATGACC 385
QY 985 GTCACTGAGAGGGGCGACCGAGCTACTGCTGCGACAGGAGTATATTTGATGAGCAAT 1044
DB 386 GTCACTGAGAGGGGCGACCGAGCTACTGCTGCGACAGGAGTATATTTGATGAGCAAT 445
QY 1045 CTCCCTCAGTCCAGCTGTTTAGAGCTGACCAACCATCTTCTTGTATCAGGAGGAT 1104
DB 446 CTCCCTCAGTCCAGCTGTTTAGAGCTGACCAACCATCTTCTTGTATCAGGAGGAT 505
QY 1105 GACATCATCTTATTCAGTGGCAAGTTTCTGCTTGA 1143
DB 506 GACATCATCTTATTCAGTGGCAAGTTTCTGCTTGA 544

RESULT 5
US-10-029-386-10599
Sequence 10599, Application US/10029386
Publication No. US20030194704A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: AEWICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 10599
LENGTH: 521
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO CHR18.3
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.92
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1
OTHER INFORMATION: EST HUMAN HIT: AW381394.1, EVALUATION 1.00e-58
OTHER INFORMATION: SWISSPROT HIT: P01014, EVALUATION 6.00e-07
OTHER INFORMATION: NT HIT: D88575.1, EVALUATION 3.00e-61
US-10-029-386-10599

Query Match 10.6%; Score 121; DB 13; Length 521;
Best Local Similarity 100.0%; Pred. No. 5e-54;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO CHR18.3
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2
OTHER INFORMATION: SWISSPROT HIT: 002739, EVALUATION 5.00e-24
OTHER INFORMATION: NT HIT: D88575.1, EVALUATION 0.00e+00
OTHER INFORMATION: EST HUMAN HIT: B1871335.1, EVALUATION 0.00e+00
US-10-029-386-2145

Query Match 34.9%; Score 399; DB 13; Length 399;
Best Local Similarity 100.0%; Pred. No. 2.9e-203;
Matches 399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 745 ATTGAAACAAACTGACCTTTTCAGATCTTAATGATGAGCAATCCAGGCGAATGACC 804
DB 1 ATTGAAACAAACTGACCTTTTCAGATCTTAATGATGAGCAATCCAGGCGAATGACC 60
QY 805 TCTAAGTATGTTGAGGTATTTTTCCTCAGTTCAGATGAGCAATCCAGGCGAATGACC 864
DB 61 TCTAAGTATGTTGAGGTATTTTTCCTCAGTTCAGATGAGCAATCCAGGCGAATGACC 120
QY 865 CAATATTGAGAGCCCTAGGGCTGAAAGATATCTTTGATGATCCAAAGCAGATCTCT 924
DB 121 CAATATTGAGAGCCCTAGGGCTGAAAGATATCTTTGATGATCCAAAGCAGATCTCT 180
QY 925 GGGATTGCTCGGGGGTCTCTGTATATATCAAGATGATGAGCAATCCAGGCGAATGACC 984
DB 181 GGGATTGCTCGGGGGTCTCTGTATATATCAAGATGATGAGCAATCCAGGCGAATGACC 240
QY 985 GTCACTGAGAGGGGCGACCGAGCTACTGCTGCGACAGGAGTATATTTGATGAGCAAT 1044
DB 241 GTCACTGAGAGGGGCGACCGAGCTACTGCTGCGACAGGAGTATATTTGATGAGCAAT 300
QY 1045 CTCCCTCAGTCCAGCTGTTTAGAGCTGACCAACCATCTTCTTGTATCAGGAGGAT 1104
DB 301 CTCCCTCAGTCCAGCTGTTTAGAGCTGACCAACCATCTTCTTGTATCAGGAGGAT 360
QY 1105 GACATCATCTTATTCAGTGGCAAGTTTCTGCTTGA 1143
DB 361 GACATCATCTTATTCAGTGGCAAGTTTCTGCTTGA 399

RESULT 4
US-10-029-386-7445
Sequence 7445, Application US/10029386
Publication No. US20030194704A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: AEWICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 7445
LENGTH: 592
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO CHR18.3
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6

QY 335 AGGACTACATGAGTCCGGAATAATTATACGATGCCAAAGTGGAGCGAGTGCATT 394
Db 326 AGGACTACATGAGTCCGGAATAATTATACGATGCCAAAGTGGAGCGAGTGCATT 385
QY 395 CGATCATTTAGAGAGCACTAGAGTAAATTAATAAGTGGTGGAAATGAACACATG 454
Db 386 CGATCATTTAGAGAGCACTAGAGTAAATTAATAAGTGGTGGAAATGAACACATG 445
QY 455 G 455
Db 446 G 446

RESULT 6
US-09-736-457-233/c
; Sequence 233, Application US/09736457
; Patent No. US20020168637A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mammion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1964
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 233
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-233

Query Match 10.6%; Score 121; DB 10; Length 600;
Best Local Similarity 100.0%; Pred. No. 5e-54; Indels 0; Gaps 0;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1023 AGTAAATATTGTAGAAAGCAACTCCCTCAGTCCAGCGTGTAGAGCTGACCCCAATT 1082
Db 500 AGTAAATATTGTAGAAAGCAACTCCCTCAGTCCAGCGTGTAGAGCTGACCCCAATT 541
QY 1083 CCTATTGTTATCAGGAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGCCCTTG 1142
Db 540 CCTATTGTTATCAGGAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGCCCTTG 481
QY 1143 A 1143
Db 480 A 480

RESULT 7
US-09-902-941-233/c
; Sequence 233, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 233
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-902-941-233

Query Match 10.6%; Score 121; DB 10; Length 600;
Best Local Similarity 100.0%; Pred. No. 5e-54; Indels 0; Gaps 0;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1023 AGTAAATATTGTAGAAAGCAACTCCCTCAGTCCAGCGTGTAGAGCTGACCCCAATT 1082
Db 500 AGTAAATATTGTAGAAAGCAACTCCCTCAGTCCAGCGTGTAGAGCTGACCCCAATT 541
QY 1083 CCTATTGTTATCAGGAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGCCCTTG 1142
Db 540 CCTATTGTTATCAGGAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGCCCTTG 481
QY 1143 A 1143
Db 480 A 480

RESULT 8
US-09-849-626-233/c
; Sequence 233, Application US/09849626
; Publication No. US20020197669A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya
; APPLICANT: Fanger, Gary
; APPLICANT: Wang, Aijun
; APPLICANT: Wang, Tongtong
; APPLICANT: Switzer, Anne
; APPLICANT: McNeill, Patricia
; APPLICANT: Clapper, Jonathan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C16
; CURRENT APPLICATION NUMBER: US/09/849,626
; CURRENT FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 1926
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 233
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-849-626-233

Query Match 10.6%; Score 121; DB 10; Length 600;
Best Local Similarity 100.0%; Pred. No. 5e-54; Indels 0; Gaps 0;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1023 AGTAAATATTGTAGAAAGCAACTCCCTCAGTCCAGCGTGTAGAGCTGACCCCAATT 1082
Db 500 AGTAAATATTGTAGAAAGCAACTCCCTCAGTCCAGCGTGTAGAGCTGACCCCAATT 541
QY 1083 CCTATTGTTATCAGGAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGCCCTTG 1142
Db 540 CCTATTGTTATCAGGAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGCCCTTG 481
QY 1143 A 1143
Db 480 A 480

RESULT 9

US-09-476-300-233/c
; Sequence 233, Application US/09476300
; Publication No. US20030125245A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C3
; CURRENT APPLICATION NUMBER: US/09/476,300
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 785
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 233
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-476-300-233

Query Match 10.6%; Score 121; DB 11; Length 600;
Best Local Similarity 100.0%; Pred. No. 5e-54;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1023 AAGTAATATTGTAGAAAAGCAACTCCCTCAGTCCACGCTGTTTAGAGCTGACACCCCAATT 1082
DB 600 AAGTAATATTGTAGAAAAGCAACTCCCTCAGTCCACGCTGTTTAGAGCTGACACCCCAATT 541
QY 1083 CCTATTGTTATCAGGAGGATGACATCATCTTATTTCAGTGGCAAGTTTCTTGCCCTTG 1142
DB 540 CCTATTGTTATCAGGAGGATGACATCATCTTATTTCAGTGGCAAGTTTCTTGCCCTTG 481

QY 1143 A 1143

DB 480 A 480

RESULT 10

US-10-113-872-233/c
; Sequence 233, Application US/10113872
; Publication No. US20030170255A1
; GENERAL INFORMATION:
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C19
; CURRENT APPLICATION NUMBER: US/10/113,872
; CURRENT FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 233
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-113-872-233

Query Match 10.6%; Score 121; DB 13; Length 600;
Best Local Similarity 100.0%; Pred. No. 5e-54;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1023 AAGTAATATTGTAGAAAAGCAACTCCCTCAGTCCACGCTGTTTAGAGCTGACACCCCAATT 1082
DB 600 AAGTAATATTGTAGAAAAGCAACTCCCTCAGTCCACGCTGTTTAGAGCTGACACCCCAATT 541
QY 1083 CCTATTGTTATCAGGAGGATGACATCATCTTATTTCAGTGGCAAGTTTCTTGCCCTTG 1142
DB 540 CCTATTGTTATCAGGAGGATGACATCATCTTATTTCAGTGGCAAGTTTCTTGCCCTTG 481

QY 1143 A 1143

DB 480 A 480

RESULT 11

US-10-017-754-233/c
; Sequence 233, Application US/10017754
; Publication No. US20030054363A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C18
; CURRENT APPLICATION NUMBER: US/10/017,754
; CURRENT FILING DATE: 2001-10-29
; NUMBER OF SEQ ID NOS: 2004
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 233
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-017-754-233

Query Match 10.6%; Score 121; DB 15; Length 600;
Best Local Similarity 100.0%; Pred. No. 5e-54;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1023 AAGTAATATTGTAGAAAAGCAACTCCCTCAGTCCACGCTGTTTAGAGCTGACACCCCAATT 1082
DB 600 AAGTAATATTGTAGAAAAGCAACTCCCTCAGTCCACGCTGTTTAGAGCTGACACCCCAATT 541
QY 1083 CCTATTGTTATCAGGAGGATGACATCATCTTATTTCAGTGGCAAGTTTCTTGCCCTTG 1142
DB 540 CCTATTGTTATCAGGAGGATGACATCATCTTATTTCAGTGGCAAGTTTCTTGCCCTTG 481

QY 1143 A 1143

DB 480 A 480

RESULT 12

US-10-029-386-24299
; Sequence 24299, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL F
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: ABOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annotmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 24299
; LENGTH: 119
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR18.3

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.92
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1
OTHER INFORMATION: SWISSPROT HIT: P29508, EVALUE 2.00e-05
OTHER INFORMATION: SKI-SPROT HIT: A931394.1, EVALUE 3.00e-58
OTHER INFORMATION: NT HIT: D88575.1, EVALUE 8.00e-61
US-10-029-386-24299

Query Match 10.4%; Score 119; DB 13; Length 119;
Best Local Similarity 100.0%; Pred. No. 5.2e-53;
Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 337 GACTACATTGAGTGTCCGAAAAAATTATACGATGCCAAAGTGGAGCGAGTTGACTTTACG 396
DB 1 GACTACATTGAGTGTCCGAAAAAATTATACGATGCCAAAGTGGAGCGAGTTGACTTTACG 60
QY 397 AATCATTTAGAGACACTAGACGTATATTAACTGGGTTGAAATGAACACATG 455
DB 61 AATCATTTAGAGACACTAGACGTATATTAACTGGGTTGAAATGAACACATG 119

RESULT 13
US-09-140-719-32
Sequence 32, Application US/09140719
Patent No. US20010026931A1
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Ruyuki
APPLICANT: TSUROOKA, No. US20010026931A1
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. US20010026931A1
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. US20010026931A1
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-140-719-32

Query Match 2.2%; Score 25; DB 9; Length 37;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTCCCTTGCTGCGCAATG 25
DB 13 ATGGCTCCCTTGCTGCGCAATG 37

RESULT 14
US-10-091-442-32
Sequence 32, Application US/10091442
Publication No. US20020164711A1
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Ruyuki
APPLICANT: TSUROOKA, No. US20020164711A1
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. US20020164711A1
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. US20020164711A1
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/091,442
FILING DATE: 07-Mar-2002
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

Mon Dec 15 08:51:17 2003

TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 32:
US-10-091-442-32

Query Match 2.2%; Score 25; DB 14; Length 37;
Best Local Similarity 100.0%; Pred. No. 0.04;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCCTCCCTTCTGCGAGCAATG 25
DB 13 ATGGCCTCCCTTCTGCGAGCAATG 37

RESULT 15
US-09-140-719-14
; Sequence 14, Application US/09140719
; Patent No. US20010026931A1
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Fuyuki
; APPLICANT: TSURUOKA, No. US20010026931A1no
; APPLICANT: NAKAZATO, Hiroshi
; APPLICANT: MIURA, Kenju
; APPLICANT: ISHIDA, No. US20010026931A1unhiro
; APPLICANT: KURIHARA, Tatsuya
; APPLICANT: YAMAICHI, Kozo
; APPLICANT: YAMAGUCHI, No. US20010026931A1omi
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/140,719
; FILING DATE: 08-AUG-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/474,661
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-067339
; FILING DATE: 04-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-247
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-140-719-14

Query Match 2.0%; Score 23; DB 9; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 791 CAAGGCGAATGACCTCTAAGTAT 813
DB 1 CAAGGCGAATGACCTCTAAGTAT 23

Search completed: December 14, 2003, 04:42:01
Job time : 430 secs

Result No.	Score	Query			ID	Description
		Match	Length	DB		
c	555	48.6	638	9	AW381394	PMO-H7030
	536	46.9	824	12	BI1871335	BI1871335
	510	44.6	1026	10	BG682538	60264811
	447	39.1	1098	13	BX284040	BX284040

Seq primer: puc 18 forward
High quality sequence start: 12
High quality sequence stop: 636.
Location/Qualifiers

FEATURES

source
1. .638
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_stage="Adult"
/clone_lib="HT0302"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
175 a 149 c 132 g 182 t

BASE COUNT 175 a 149 c 132 g 182 t
ORIGIN
Query Match 48.6%; Score 555; DB 9; Length 638;
Best Local Similarity 99.8%; Pred. No. 2.5e-286;
Matches 605; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 32 TTGCTTCAACCTGTTTCAGAGATGGATGACATCAAGGAATCGAAATGTTCTTTT 91
Db 608 TTGCTTCAACCTGTTTCAGAGATGGATGACATCAAGGAATCGAAATGTTCTTTT 549
QY 92 CTTCTGAGCTCTTCGTCGCCCTGCGCTCGCTTGGCGCTCAGATGACTCCC 151
Db 548 CTTCTGAGCTCTTCGTCGCCCTGCGCTCGCTTGGCGCTCAGATGACTCCC 489
QY 152 TCTCCAGATGATAGTTGCTTATGTTTAACTCTCAGGATATGGAATCTTTCTA 211
Db 488 TCTCCAGATGATAGTTGCTTATGTTTAACTCTCAGGATATGGAATCTTTCTA 429
QY 212 ATAGTCAGTCAGGCTCCAGTCTCAACTGAAAGAGTTTCTTCTGATATTAATGATCCC 271
Db 428 ATAGTCAGTCAGGCTCCAGTCTCAACTGAAAGAGTTTCTTCTGATATTAATGATCCC 369
QY 272 ACAGGATATGATCTCAGATGTTGATGGCTTTTCTGCTCAAAAGCTGATGGCTTC 331
Db 368 ACAGGATATGATCTCAGATGTTGATGGCTTTTCTGCTCAAAAGCTGATGGCTTC 309
QY 332 ATAGGACTATGATGTTGCTGCGGAAATATACGATCCAAAGTGGAGCGAGTTGACT 391
Db 308 ATAGGACTATGATGTTGCTGCGGAAATATACGATCCAAAGTGGAGCGAGTTGACT 249
QY 392 TTACGAATCATTTAGAGACACATGACGATTAATTAATAGTGGTTGAAATGAAACAC 451
Db 248 TTACGAATCATTTAGAGACACATGACGATTAATTAATAGTGGTTGAAATGAAACAC 189
QY 452 ATGGCAAAATCAGAACGATGTTGAGTGGGATAGCTCATCTCTGTAATGGTGC 511
Db 188 ATGGCAAAATCAGAACGATGTTGAGTGGGATAGCTCATCTCTGTAATGGTGC 129
QY 512 TGGTGAATGCTGTACTTCAAAAGGCAAGTGGCAATCAGCTTCCAAAGAGGAAACCA 571
Db 128 TGGTGAATGCTGTACTTCAAAAGGCAAGTGGCAATCAGCTTCCAAAGAGGAAACCA 69
QY 572 TAAATGGCAATTCAAATCTCCCAAGTGTCTGGGAGGCGAGTGGCGATGATGATCAGG 631
Db 66 TAAATGGCAATTCAAATCTCCCAAGTGTCTGGGAGGCGAGTGGCGATGATGATCAGG 9
QY 632 AACGGA 637
Db 8 AACGGA 3

RESULT 2
BI871335 824 bp mRNA linear EST 11-OCT-2001
LOCUS 603392761F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5402966 5',
DEFINITION

mRNA sequence.

BI871335
BI871335.1 GI:16045010
EST.
Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.mci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12027 row: i column: 15
High quality sequence stop: 824.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1. 824
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5402966"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 90"
/note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH MGC Library."
232 a 186 c 164 g 242 t

BASE COUNT

ORIGIN

Query Match 46.9%; Score 536; DB 12; Length 824;
Best Local Similarity 100.0%; Pred. No. 4.5e-276;
Matches 536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 608 AGGCACTGCCATGATGCTCAGAACCGAGTCTCAATTTGCTGTTATGAGGACCCAT 667
Db 12 AGGCACTGCCATGATGCTCAGAACCGAGTCTCAATTTGCTGTTATGAGGACCCAT 71
QY 668 CAATGAAGATTTCTTGAGCTCAGATACATGCTGGCATAAATGATGCTTCTGCTGCTG 727
Db 72 CAATGAAGATTTCTTGAGCTCAGATACATGCTGGCATAAATGATGCTTCTGCTGCTG 131
QY 728 AGAATGACCTCTCTGAATTAATAAACAACATGACCTTTCAGATCTTAATGAATGGACCA 787
Db 132 AGAATGACCTCTCTGAATTAATAAACAACATGACCTTTCAGATCTTAATGAATGGACCA 191
QY 788 ATCCCAAGGCGATGACCTCTAAGTATGTTGAGTATTTTTCCTCAGTTCAAGATAGAGA 847
Db 192 ATCCCAAGGCGATGACCTCTAAGTATGTTGAGTATTTTTCCTCAGTTCAAGATAGAGA 251
QY 848 AGAATTAATGAATTAATAAACAACATGACCTTTCAGATCTTAATGAATGGACCA 907
Db 252 AGAATTAATGAATTAATAAACAACATGACCTTTCAGATCTTAATGAATGGACCA 311
QY 908 CCAAGAGCAGATCTCTCGGATGCTTCGGGGGCTCGCTCTGTATATATCAGGATGATGC 967
Db 312 CCAAGAGCAGATCTCTCGGATGCTTCGGGGGCTCGCTCTGTATATATCAGGATGATGC 371
QY 968 ACAATCTTACATAGAGGTCTCAGGAGGCGACCGAGGCTTACTGCTGCCACAGAGTA 1027
Db 372 ACAATCTTACATAGAGGTCTCAGGAGGCGACCGAGGCTTACTGCTGCCACAGAGTA 431
QY 1028 ATATTGTAGAAAGCAACTCCCTCAGTCCACGCTGTTTAGAGTGTGACCACTTCCTAT 1087

```

Db 432 ATATGTAGAAAGCACTCCCTCAGTCACGCTGTTAGAGCTGACACCACTTCCTAT 491
QY 1088 TTGTATCAGGAAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGCCCTTGA 1143
Db 452 TTGTATCAGGAAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGCCCTTGA 547

RESULT 3
EG682538 1026 bp mRNA linear EST 01-MAY-2001
LOCUS 602624811F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4749699 5'
DEFINITION mRNA sequence.
ACCESSION EG682538
VERSION 1 (bases 1 to 1026)
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1026)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10603 row: f column: 04
High quality sequence start: 3
High quality sequence stop: 791.

FEATURES
source
1..1026
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4749699"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Skn4"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: oligo dr.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 303 a 221 c 213 g 289 t
ORIGIN
Query Match 44.6%; Score 510; DB 10; Length 1026;
Best Local Similarity 99.8%; Pred. No. 4.5e-262;
Matches 560; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 583 TTCAAATCTCCAGTGTCTCGGAAGGACATGCCATCATGATCAGCAAGCAAGATTC 642
Db 9 TTCAAATCTCCAGTGTCTCGGAAGGACATGCCATCATGATCAGCAAGCAAGATTC 68

QY 643 AATTTGCTCTGTTATGAGGACCCATCAATGAAGATTTCTTGAGCTCAGATACAAATGTTGGC 702
Db 69 AATTTGCTCTGTTATGAGGACCCATCAATGAAGATTTCTTGAGCTCAGATACAAATGTTGGC 128

QY 703 ATAAACATGATGATGCTGCTGCTGAGATGACCTCTCTGAAATTTGAACAAACTGACC 762
Db 129 ATAAACATGATGATGCTGCTGCTGAGATGACCTCTCTGAAATTTGAACAAACTGACC 188

QY 763 TTTGAGATCTAATGATGGAACCAATCCAGGCAATGACCTCTAGTATGTTGAGSTA 822
Db 189 TTTGAGATCTAATGATGGAACCAATCCAGGCAATGACCTCTAGTATGTTGAGSTA 248

QY 823 TTTTTCCTCAGTTCAGATAGACAGAAATTATGAAGTAAAGAAACAATATTGAGAGCCCTA 882

```

```

Db 249 TTTTTCCTCAGTTCAGATAGACAGAAATTATGAATGAAGAAACAATATTGAGAGCCCTA 308
QY 883 GGGCTGAAGATATCTTTTATGATGAATCCAAAGCAGATCTCTCTGGGATTCGTTGGGGGGT 942
Db 309 GGGCTGAAGATATCTTTTATGATGAATCCAAAGCAGATCTCTCTGGGATTCGTTGGGGGT 368
QY 943 GGTCTGATATATCAAGATGATCCCAATCTTACATAGAGTTCACCTGAGGAGGACCC 1002
Db 369 GGTCTGATATATCAAGATGATCCCAATCTTACATAGAGTTCACCTGAGGAGGACCC 428
QY 1003 GAGGCTACTCTCTCCACAGGAGTAATATTGTAGAAAAGCACTCCCTCAGTCCACGGCTG 1062
Db 429 GAGGCTACTCTCTCCACAGGAGTAATATTGTAGAAAAGCACTCCCTCAGTCCACGGCTG 488
QY 1063 TTTAGAGTCAACCACTTCTTATTTGTATCAGGAAGGATGATCATCTATTATTCAGT 1122
Db 489 TTTAGAGTCAACCACTTCTTATTTGTATCAGGAAGGATGATCATCTATTATTCAGT 548
QY 1123 GGCAGAGTTCTTGCCCTTGA 1143
Db 549 GGCAGAGTTCTTGCCCTTGA 569

RESULT 4
EX284040 498 bp mRNA linear EST 05-MAR-2003
LOCUS EX284040 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4749699 5'
DEFINITION IMAGE:4749699, mRNA sequence.
ACCESSION EX284040
VERSION EX284040.1 GI:28848494
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 498)
AUTHORS Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,B., Peters,M.,
Radeiof,U., Schneider,D. and Korn,B.
Human UnigeneSet - RZPD3
Unpublished
COMMENT Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD: IMAGp998F0410603.
RZPDLIB: I.M.A.G.E. cDNA Clone Collection;
Human UnigeneSet - RZPD3 (RZPDLIB No.972)
http://www.rzpd.de/CloneCards/cgi-
bin/showlib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
SP6, Primer sequence: ATTTAGGTGACATAATAG.

FEATURES
source
1..498
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGp998F0410603; IMAGE:4749699"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Skn4"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: oligo dr.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 153 a 101 c 114 g 130 t
ORIGIN

```


VERSION	RG319597.1	GI:13129272
KEYWORDS	Est.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1 [bases 1 to 405]	
AUTHORS	Schlager, J. J., Benjamin, H. R., Ali, K., Levine, C. F., Dodds, A. D., Avery, D. P., Clark, J. H. and Hofmann, C. M.	
TITLE	Human Spidermal Keratinocyte Subtraction Library-Downregulated Transcripts	
JOURNAL	Unpublished	

```

COMMENT
Contact: Schlager, J.J.
Molecular Toxicology
United States Army Medical Research Institute of Chemical Defense
(USAMRICD)
3100 Ricketts Point Road; ATTN: MCMR-UV-PA; AFG-EA, MD 21010-5400,
USA
Tel: 410 436 1940
Fax: 410 436 1960
Email: John.Schlager@MEDD.ARMY.MIL.
Location/Qualifiers
1..405
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/sex="female"
/tissue_type="skin"
/cell_type="primary keratinocyte"
/cell_line="Clonetics 4075"
/dev stage="adult"

FEATURES
source

```

```

/lab host="E. coli (ampicillin-resistant)"
/clone.lib=Human Keratinocyte Subtraction Library-
Downregulated Transcripts"
/notes=vector: pT-Adv; Library preparation: Sequence
isolated using PCR-select (Clontech) subtraction library
construction (Diachenko et al., (1996) Proc. Natl. Acad
Sci. USA 93: 6025-6030) after sulfur mustard exposure of
primary human adult epidermal keratinocytes. Subtraction
library cloned for random sequence selection into pT-Adv
vector. "
BASE COUNT      124 a 81 c 90 g 110 t
ORIGIN
Query Match      35.4%; Score 405; DB 10; Length 405;
Best Local Similarity 100.0%; Pred.No. 8.4e-206;
Matches 405; Conservative 0; Mismatches 0; Gaps 0
Qy 714 CGTTCTGCTGCTGAGATGACCTCTCTGAAATGGAAACAAACTGACCTTTTCAGAACT 773
Db 1 CGTTCTGCTGCTGAGATGACCTCTCTGAAATGGAAACAAACTGACCTTTTCAGAACT 60
Qy 774 ANTGAATGACCAATCCAAAGCGGATGACCTCTTAAGTATGTTGAGGTATTTTTTCCTCA 833
Db 61 AATGGAAATGGACCAATCCAAAGCGGAAATGACCTCTTAAGTATGTTGAGGTATTTTTTCCTCA 120
Qy 834 GTTCAAGATAGAGAAGAATTTAGAAATGAAACAATAATTTGAGAGCCCTAGGGGCTGAAAGA 893
Db 121 GTTCACATAGAGAAGATTTTGAANAGAACAAATATTTGAGAGCCCTAGGGCTGAAGA 180
Qy 894 TATCTTTGATGATCCAAAGCAGATCTCTCTGGGATGCTTCGCGGGGGTGCCTGTATAT 953
Db 181 TATCTTTGATGATCCAAAGCAGATCTCTCTGGGATGCTTCGCGGGGGTGCCTGTATAT 240

```

Qy	954	ATCAGGATGATGCACAAATCTTATCATAGAGGTCACTGAGGAGGGCACCGAGGCTACTGC	1010
Db	241	ATCAGGATGATGCACAAATCTTATCATAGAGGTCACTGAGGAGGGCACCGAGGCTACTGC	300
Qy	1014	TGCCACAGGAGTAATATGTTAGAAAACCACTCCCTCAGTCCACGGCTGTTTAGAGCTGA	1070
Db	301	TGCCACAGGAGTAATATGTTAGAAAACCACTCCCTCAGTCCACGGCTGTTTAGAGCTGA	360
Qy	1074	CCACCCATCTCTATTGTTTATCAGGAAGGATGACATCATCTTATT	1116

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4069704"
/tissue_type="hybernephroma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NH_MGC_58"
/notes="Organ: kidney; Vector: pDNR-LIB (Clontech); Site_1: 5'f11 (ggcgcttcggcc); Site_2: 5'f11 (ggccattatggcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-ATCTAGAGCGCGGCGGCACATG-dn(30)BN-3' sequence: 5'-ATCTAGAGCGCGGCGGCACATG-dn(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.35 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."
```

BASE COUNT	210 a	129 c	149 g	139 t	
ORIGIN					
	Query Match	15.8%;	Score 181;	DB 10;	Length 627;
	Best Local Similarity	100.0%;	Pred.No. 2.2e-85;		
	Matches 181;	Conservative	0;	Mismatches	0;
				Indels	0;
Qy	426	TAATAGTGGGTTGAAATGAAACACATGCAATATCAAGACGTCGATTTGGTGAAGTGG	485		
Db	1	TAATAAGTGGGTTGAAATGAAACACATGCAATATCAAGACGTCGATTTGGTGAAGTGG	60		
Qy	486	CATAAGCTCATCTGTCGTAAATGGTGGTGGTGAATGCTGTACTTCAAAAGCAAGTGGCA	545		
Db	61	CATAAGCTCATCTGTCGTAAATGGTGGTGGTGAATGCTGTACTTCAAAAGCAAGTGGCA	120		
Qy	546	ATCAGGCTTCACAGAGGGAACCATTAATTTGCCATTTCAATCTCCCAAGTCTCTGG	605		
Db	121	ATCAGGCTTCACAGAGGGAACCATTAATTTGCCATTTCAATCTCCCAAGTCTCTGG	180		
Qy	606	G	606	G	606
Db	181	G	181	G	181


```

/clone="hbc2326"
/clone.lib="Human pancreatic islet"
/notes="Vector: Lambda ZAPII; Site 1: Eco RI; Site 2: Xho I; mRNA was prepared from normal adult human islets. cDNA was directionally synthesized from the Xho I in the vector to the EcoRI site. cDNA was size fractionated to remove sequences <1000 bp in size."
BASE COUNT      104 a   78 c   82 g   89 t   10 others
ORIGIN
Query Match      6.3%; Score 76; DB 14; Length 363;
Best Local Similarity 100.0%; Pred. No. 4.7e-29; Mismatches 0; Indels 0; Gaps 0;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 671 TGAAGATCTTCTGAGCTCAGATACAAATGGTGGCATAAACATGTACGTTCTGCTGCTGAGA 730
      |||||
Db 180 TGAAGATCTTCTGAGCTCAGATACAAATGGTGGCATAAACATGTACGTTCTGCTGCTGAGA 239

QY 731 ATGACCTCTCTGAAT 746
      |||||
Db 240 ATGACCTCTCTGAAT 255

RESULT 13
BF243008
LOCUS      1640 bp mRNA linear EST 14-NOV-2000
DEFINITION mRNA sequence.
ACCESSION BF243008
VERSION    BF243008.1 GI:11156936
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 1640)
AUTHORS   NIH-MGC http://mgi.nci.nih.gov/
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT    Contact: Robert Strausberg, Ph.D.
           Email: cgabbs-remail.nih.gov
           Tissue Procurement: ATCC
           CDNA Library Preparation: CLONETECH Laboratories, Inc.
           CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Incyte Genomics, Inc.
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           plate: LLCM985 row: 1 column: 24
           High quality sequence stop: 503.
           Location/Qualifiers
             1. .1640
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:4106207"
               /tissue_type="from acute myelogenous leukemia"
               /lab_host="DH10B (T1 phage-resistant)"
               /clone_lib="NIH MGC 55"
               /notes="Organ: bone marrow; Vector: pDNR-LIB (Clontech);
               Site 1: SfiI (ggcgctcgcc); Site 2: SfiI (ggcgctcgcc)
               ); Double-stranded cDNA was prepared from cell line RNA.
               5 and 3' adaptors were used in cloning as follows: 5'
               adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor
               sequence: 5'-ATTCTAGAGCGGCGGCGGCGATG-dt(30)BN-3'
               (where B = A, C, or G and N = A, C, G, or T). Average
               insert size 1.65 kb (range 0.9-4.0 kb). 14/15 colonies
               contained inserts by PCR. This library was enriched for
               full-length clones and was constructed by Clontech
               Laboratories (Palo Alto, CA)."
BASE COUNT      543 a   377 c   418 g   302 t
ORIGIN

```

```

Query Match      6.3%; Score 72; DB 10; Length 1640;
Best Local Similarity 100.0%; Pred. No. 1.2e-26; Mismatches 0; Indels 0; Gaps 0;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1072 GACCACCCATCTCTATTGTTATCAGGAAGGATGACATCATCTTATTCAGTGGCAAGTT 1131
      |||||
Db 1 GACCACCCATCTCTATTGTTATCAGGAAGGATGACATCATCTTATTCAGTGGCAAGTT 60

QY 1132 TCTTGGCCCTGA 1143
      |||||
Db 61 TCTTGGCCCTGA 72

RESULT 14
AK014524
LOCUS      1889 bp mRNA linear HTC 05-DEC-2002
DEFINITION Mus musculus 0 day neonate skin cDNA, RIKEN full-length enriched
library, clone:4631416M05 product:serine (or cysteine) proteinase
inhibitor, clone B (ovalbumin), member 7, full insert sequence.
ACCESSION AK014524
VERSION    AK014524.1 GI:12852434
KEYWORDS  HTC; CAP trapper.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1
AUTHORS   Carninci, P. and Hayashizaki, Y.
TITLE     High-efficiency full-length cDNA cloning
JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
MEDLINE   99279253
PubMed    10349636

REFERENCE  2
AUTHORS   Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE   20493374
PubMed    11042159

REFERENCE  3
AUTHORS   Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Kono, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamanoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanuki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, I., Matsuda, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE   20530913
PubMed    11076861

REFERENCE  4
AUTHORS   Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
Saito, T., Okazaki, Y., Gojohori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischman, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
Quackenbush, J., Schriml, L. M., Staab, F., Suzuki, M., Tomita, M.,
Wagner, L., Wastio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,
Fleischer, C., Fujita, M., Kamiya, M., Lee, N. H., Lyons, P.,
Machinomi, L., Mashima, J., Mazzarelli, J., Mombaeerts, P., Nordone, P.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seva, T., Shibata, Y., Storch, K. F., Suzuki, H.,
Tayo-Oka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohzuki, S.
and Hayashizaki, Y.

```


electrophoresis. Vector DNA was prepared from a derivative of pMD42 (G14732114|9b|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 94 a 64 c 75 g 80 t
ORIGIN

Query Match 2.5%; Score 29; DB 28; Length 113;
Best Local Similarity 100.0%; Pred. No. 0.0008;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTCCCTTGCTGCAGCAATGCAGA 29
|||
Db 250 ATGGCTCCCTTGCTGCAGCAATGCAGA 222
|||

Search completed: December 14, 2003, 06:48:14
Job time : 2784 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: December 13, 2003, 23:52:22 ; Search time 4370 Seconds
(without alignments)
10700.161 Million cell updates/sec
Title: US-09-508-997a-1
Perfect score: 1143
Sequence: 1 agggcctccctgtgcagc.....gcaaggttttgccttga 1143
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 2888711 seqs, 20454813386 residues
Total number of hits satisfying chosen parameters: 5777422
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match C%
Maximum Match 100%
Listing first 45 summaries

Database :					GenEmbl :				
1:	gb	ba	*		1:	gb	ba	*	
2:	gb	htg	*		2:	gb	htg	*	
3:	gb	in	*		3:	gb	in	*	
4:	gb	om	*		4:	gb	om	*	
5:	gb	ov	*		5:	gb	ov	*	
6:	gb	pat	*		6:	gb	pat	*	
7:	gb	ph	*		7:	gb	ph	*	
8:	gb	pl	*		8:	gb	pl	*	
9:	gb	pr	*		9:	gb	pr	*	
10:	gb	ro	*		10:	gb	ro	*	
11:	gb	sts	*		11:	gb	sts	*	
12:	gb	sy	*		12:	gb	sy	*	
13:	gb	un	*		13:	gb	un	*	
14:	gb	vi	*		14:	gb	vi	*	
15:	em	ba	*		15:	em	ba	*	
16:	em	fun	*		16:	em	fun	*	
17:	em	hum	*		17:	em	hum	*	
18:	em	in	*		18:	em	in	*	
19:	em	mu	*		19:	em	mu	*	
20:	em	om	*		20:	em	om	*	
21:	em	or	*		21:	em	or	*	
22:	em	ov	*		22:	em	ov	*	
23:	em	pat	*		23:	em	pat	*	
24:	em	ph	*		24:	em	ph	*	
25:	em	pl	*		25:	em	pl	*	
26:	em	ro	*		26:	em	ro	*	
27:	em	sts	*		27:	em	sts	*	
28:	em	un	*		28:	em	un	*	
29:	em	vi	*		29:	em	vi	*	
30:	em	htg	hum	*	30:	em	htg	hum	*
31:	em	htg	inv	*	31:	em	htg	inv	*
32:	em	htg	other	*	32:	em	htg	other	*
33:	em	htg	mus	*	33:	em	htg	mus	*
34:	em	htg	pln	*	34:	em	htg	pln	*
35:	em	htg	rod	*	35:	em	htg	rod	*
36:	em	htg	mam	*	36:	em	htg	mam	*
37:	em	htg	virt	*	37:	em	htg	virt	*
38:	em	sy	*		38:	em	sy	*	
39:	em	htgo	hum	*	39:	em	htgo	hum	*
40:	em	htgo	mus	*	40:	em	htgo	mus	*
41:	em	htgo	other	*	41:	em	htgo	other	*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result		Query		DB		ID		Description	
No.	Score	Match	Length	No.	Score	Match	Length	No.	Score
1	1143	100.0	1143	6	BD094240	Model ani		BD094240	Model ani
2	1143	100.0	1950	6	AR052445	Sequence		AR052445	Sequence
3	1143	100.0	1950	6	AR082433	Sequence		AR082433	Sequence
4	1143	100.0	1950	6	EO8396	cDNA encodi		EO8396	cDNA encodi
5	1143	100.0	1950	9	D88575	Homo sapien		D88575	Homo sapien
6	1143	100.0	2249	9	AF027866	Homo sapi		AF027866	Homo sapi
7	781.4	68.4	1386	10	AF105328	Mus muscu		AF105328	Mus muscu
8	755.8	66.1	1143	10	AF105329	Rattus no		AF105329	Rattus no
9	755.8	66.1	1229	6	BD096261	Method fo		BD096261	Method fo
10	753	65.9	1229	6	BD094241	Model ani		BD094241	Model ani
11	748.6	65.5	1147	6	BD094242	Model ani		BD094242	Model ani
12	399.4	34.9	157284	9	AC072051	Homo sapi		AC072051	Homo sapi
13	399.4	34.9	189092	2	AC027524	Homo sapi		AC027524	Homo sapi
14	304	26.6	316	6	AR270261	Sequence		AR270261	Sequence
15	279.4	24.4	121141	10	AC125314	Mus muscu		AC125314	Mus muscu
16	260.2	22.8	51602	2	AC129261	Rattus no		AC129261	Rattus no
17	260.2	22.8	199669	10	AC103453	Rattus no		AC103453	Rattus no
18	260.2	22.8	219071	2	AC133259	Rattus no		AC133259	Rattus no
19	254	22.2	254	6	AR270095	Sequence		AR270095	Sequence
20	240.6	21.0	1179	9	AF419953	Homo sapi		AF419953	Homo sapi
21	240.6	21.0	1179	9	AF419954	Homo sapi		AF419954	Homo sapi
22	240.6	21.0	1179	9	AF419955	Homo sapi		AF419955	Homo sapi
23	237.8	20.8	1200	6	AX301894	Sequence		AX301894	Sequence
24	237.4	20.8	1563	10	BC010313	Mus muscu		BC010313	Mus muscu
25	225.6	19.7	1512	6	A21239	H sapiens B		A21239	H sapiens B
26	225.6	19.7	1512	6	A21239	H sapiens B		A21239	H sapiens B
27	222	19.4	1482	6	A21240	H sapiens B		A21240	H sapiens B
28	222	19.4	1482	6	A21240	H sapiens B		A21240	H sapiens B
29	220.6	19.3	1279	6	AX677161	Sequence		AX677161	Sequence
30	220.6	19.3	1279	9	AF169949	Homo sapi		AF169949	Homo sapi
31	220.6	19.3	1308	6	AF176431	Sequence		AF176431	Sequence
32	220.6	19.3	1308	6	AR220500	Sequence		AR220500	Sequence
33	220.6	19.3	1308	6	AR255494	Sequence		AR255494	Sequence
34	220.6	19.3	1308	6	AR281064	Sequence		AR281064	Sequence
35	220.6	19.3	1308	6	AX365716	Sequence		AX365716	Sequence
36	219	19.2	1279	6	AX092547	Sequence		AX092547	Sequence
37	219	19.2	2654	9	HSP113122	Sequence		HSP113122	Sequence
38	219	19.2	3443	9	HSP113122	Sequence		HSP113122	Sequence
39	214.2	18.7	1327	4	D55670	Bos taurus		D55670	Bos taurus
40	211.6	18.5	1173	9	AY245781	Homo sapi		AY245781	Homo sapi
41	211.6	18.5	1173	9	BT006748	Homo sapi		BT006748	Homo sapi
42	211.6	18.5	1173	12	BT007953	Synthetic		BT007953	Synthetic
43	211.6	18.5	1245	6	AX030195	Sequence		AX030195	Sequence
44	211.6	18.5	1279	9	HSU19556	Human squa		HSU19556	Human squa
45	211.6	18.5	1711	6	AR019398	Sequence		AR019398	Sequence

ALIGNMENTS

RESULT 1
BD094240
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

BD094240
Model animals of nephropathy proliferating mesangial cells.
BD094240
BD094240.1 GI:22639828
WC 0124628-A/1.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1143)
Miyata, T.
Model animals of nephropathy proliferating mesangial cells
Patent: WO 0124628-A 1 12-APR-2001;

TOSHIO MIYATA, KIYOSHI KUROKAWA

COMMENT

OS Homo sapiens (human)

PN WC 0124628-A/1

PD 12-APR-2001

PR 06-OCT-2000 WO 2000P006988

PF 06-OCT-1999 JP 99P 285736

PI TOSHIO MIYATA

PC A01K67/027,A61P13/12,A61K45/00,C12N15/12,C12N15/85,G01N33/15,

PP G01N33/50

FEATURES

CC Key Location/Qualifiers

CT CDS Location/Qualifiers

FT 1..1143 (1)..(1140).

source /organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

BASE COUNT 336 a 230 c 261 g 316 t

ORIGIN

Query Match 100.0%; Score 1143; DB 6; Length 1143;
 Best Local Similarity 100.0%; Pred. No. 9.9e-312;
 Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCTCCCTGCTGCGAGCAATGCGAGATTTTGGTCAACCTGTTTCAGAGAGATGGAAT 60
 Db 1 ATGGCTCCCTGCTGCGAGCAATGCGAGATTTTGGTCAACCTGTTTCAGAGAGATGGAAT 60

Qy 61 GACAATCAAGGAATGGAATGTGTTCTTCTCTGAGGCTCTTCGCTGCCCTGGCC 120
 Db 61 GACAATCAAGGAATGGAATGTGTTCTTCTCTGAGGCTCTTCGCTGCCCTGGCC 120

Qy 121 CTGGTCGGCTGGGGCTCAAGATGACCTCTCTCAGATTGATAGTTGCTTCATGTT 180
 Db 121 CTGGTCGGCTGGGGCTCAAGATGACCTCTCTCAGATTGATAGTTGCTTCATGTT 180

Qy 181 AACACTGCTCAGGATGGAATCTTCTAATAGTCAGTCAGGCTCCAGTCTCAACTG 240
 Db 181 AACACTGCTCAGGATGGAATCTTCTAATAGTCAGTCAGGCTCCAGTCTCAACTG 240

Qy 241 AAAGAGTTTTTCTGATATAAATGCAATCCCAAGGATTAAGATCTCAGCATTTGGAAT 300
 Db 241 AAAGAGTTTTTCTGATATAAATGCAATCCCAAGGATTAAGATCTCAGCATTTGGAAT 300

Qy 301 GGGCTTTTCTGAAAAGTGTGCTTTTATAGACTACATTTAGCTGTGGCGAAAAA 360
 Db 301 GGGCTTTTCTGAAAAGTGTGCTTTTATAGACTACATTTAGCTGTGGCGAAAAA 360

Qy 361 TTATACGTCGCAAGTGGAGGAGTGTGCTTTTACGATCATTTAGAGACACTAGACCT 420
 Db 361 TTATACGTCGCAAGTGGAGGAGTGTGCTTTTACGATCATTTAGAGACACTAGACCT 420

Qy 421 AATAATTAATGCTGGTTGAAATGAAAACATGGCAAAATCAAGAACGTAATGTGTGAA 480
 Db 421 AATAATTAATGCTGGTTGAAATGAAAACATGGCAAAATCAAGAACGTAATGTGTGAA 480

Qy 481 GGTGGCTAAGCTCATCTGCTGTAATGCTGTGTGCTGTGCTGTCTCAAGAGCAAG 540
 Db 481 GGTGGCTAAGCTCATCTGCTGTAATGCTGTGTGCTGTGCTGTCTCAAGAGCAAG 540

Qy 541 TGGCAATCAGCTTCAACAGAGCGAAACATAAATTTGCCATTTCAATTTCCCAAGTGC 600
 Db 541 TGGCAATCAGCTTCAACAGAGCGAAACATAAATTTGCCATTTCAATTTCCCAAGTGC 600

Qy 601 TCTGGGAAGCAGTGGCCATGATGCATCAGGAACGGAAGTTCAATTTGCTGTTATGAG 660
 Db 601 TCTGGGAAGCAGTGGCCATGATGCATCAGGAACGGAAGTTCAATTTGCTGTTATGAG 660

Qy 661 GACCCATCAATGAGATTTTCTGAGCTCAGATCAATGTTGGCATTAACATGAGCTCTG 720
 Db 661 GACCCATCAATGAGATTTTCTGAGCTCAGATCAATGTTGGCATTAACATGAGCTCTG 720

Qy 721 CTGCTGAGAAATGACCTCTCTGAAAATTTGAAAAACAACTGACCTTTTCAAGATCTAATGAA 780

Db 721 CTGCTGAGATGACCTCTCTGAAATGAAAACAACTGACCTTTCAAGATCTAATGGA 780
 Qy 781 TGGACCAATCCCAAGCGAATGACCTCTAAAGTATCTTGAGGTATTTTTTCTCAGTTCAAG 840
 Db 781 TGGACCAATCCCAAGCGAATGACCTCTAAAGTATCTTGAGGTATTTTTTCTCAGTTCAAG 840

Qy 841 ATAGAGAAATATGAAATGAAACAAATATTTGAGAGCCTTAGGGCTCAAGAGATCTTT 900
 Db 841 ATAGAGAAATATGAAATGAAACAAATATTTGAGAGCCTTAGGGCTCAAGAGATCTTT 900

Qy 901 GATGATCCAAAGCAGATCTCTCGGATTTGCTTCGGGGGGTCTCTGTATATATCAAGG 960
 Db 901 GATGATCCAAAGCAGATCTCTCTCGGATTTGCTTCGGGGGGTCTCTGTATATCAAGG 960

Qy 961 ATGATGCAAAATCTTACATAGAGTCACTGAGAGGCGACCGAGGCTACTGCTGCCACA 1020
 Db 961 ATGATGCAAAATCTTACATAGAGTCACTGAGAGGCGACCGAGGCTACTGCTGCCACA 1020

Qy 1021 GGAAGTAATTTGTAGAAAAGCACTCCCTCAGTCCACGCTGTTTAGAGCTGACACCCA 1080
 Db 1021 GGAAGTAATTTGTAGAAAAGCACTCCCTCAGTCCACGCTGTTTAGAGCTGACACCCA 1080

Qy 1081 TTCTATTTGTTATCAGGAAGATGACATCATCTTATTCAGTGGCAAAAGTTTCTTGCCCT 1140
 Db 1081 TTCTATTTGTTATCAGGAAGATGACATCATCTTATTCAGTGGCAAAAGTTTCTTGCCCT 1140

Qy 1141 TGA 1143
 Db 1141 TGA 1143

RESULT 2

AR052445

LOCUS

Sequence 30 from patent US 5831030.

DEFINITION

AR052445

ACCESSION

VERSION

AR052445.1 GI:5975909

KEYWORDS

Unknown.

SOURCE

Unknown.

ORGANISM

Unclassified.

REFERENCE

1 (bases 1 to 1950)

AUTHORS

Tsujiimoto, M., Iwasa, F., Tsuruoka, N., Nakazato, H., Miura, K.,

Ishida, N., Kurihara, T., Yamaichi, K. and Yamaguchi, N.

TITLE

Antibodies specific for megakaryocyte differentiation factor

JOURNAL

Patent: US 5831030-A 30 03-NOV-1998;

FEATURES

Location/Qualifiers

1..1950

source

/organism="unknown"

BASE COUNT 570 a 384 c 407 g 589 t

ORIGIN

Query Match

100.0%; Score 1143; DB 6; Length 1950;

Best Local Similarity 100.0%; Pred. No. 1e-311;

Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCTCCCTGCTGCGAGCAATGCGAGATTTTGGTCAACCTGTTTCAGAGAGATGGAAT 60

Db 74 ATGGCTCCCTGCTGCGAGCAATGCGAGATTTTGGTCAACCTGTTTCAGAGAGATGGAAT 133

Qy 61 GACAATCAAGGAATGGAATGTGTTCTTCTCTGAGGCTCTTCGCTGCCCTGGCC 120

Db 134 GACAATCAAGGAATGGAATGTGTTCTTCTCTGAGGCTCTTCGCTGCCCTGGCC 193

Qy 121 CTGGTCGGCTGGGGCTCAAGATGACCTCTCTCAGATTGATAGTTGCTTCATGTT 180

Db 194 CTGGTCGGCTGGGGCTCAAGATGACCTCTCTCAGATTGATAGTTGCTTCATGTT 253

Qy 181 AACACTGCTCAGGATGGAATCTTCTAATAGTCAGTCAGGCTCCAGTCTCAACTG 240

Db 254 AACACTGCTCAGGATGGAATCTTCTAATAGTCAGTCAGGCTCCAGTCTCAACTG 313

241 AAAAGAGTTTTTTCTGATATAAATGCATCCACAGGATTAATGATCTCAGCATTTGTAAT 300
Db |||||
314 ABAAGAGTTTTTTCTGATATAAATGCATCCACAGGATTAATGATCTCAGCATTTGTAAT 373
Qy GGGCTTTTCTGTAAGAAAGTGTATGCTTTTATTAAGGACTACATTTGAGTGTGCGGAAAAA 360
Db |||||
374 GGGCTTTTCTGTAAGAAAGTGTATGCTTTTATTAAGGACTACATTTGAGTGTGCGGAAAAA 433
Qy TTTATGAGTGCACAAAGTGGAGCGAGTGTGCTTTTACCAATCATTTTAAAGAGACTAGACCT 420
Db |||||
434 TTATACGATGGCAAGTGGAGCGAGTGTGCTTTTACCAATCATTTTAAAGAGACTAGACCT 493
Qy 421 AATATTAAATAGTGGTGTAAATGAACACATGCGCAAAATCAAGAACCTGATTGTGAA 480
Db |||||
494 AATATTAAATAGTGGTGTAAATGAACACATGCGCAAAATCAAGAACCTGATTGTGAA 553
Qy 481 GGTGGCATTAAGCTCATCTGCTGTAAATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 540
Db |||||
554 GGTGGCATTAAGCTCATCTGCTGTAAATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 613
Qy 541 TGGCAATCAGGCTTCCACAGAGCGAATCAATTAATTTGCCAATTTCAAAATCTCCCAAGTGC 600
Db |||||
614 TGGCAATCAGGCTTCCACAGAGCGAATCAATTAATTTGCCAATTTCAAAATCTCCCAAGTGC 673
Qy 601 TCTGGGAAGGAGTGCCTCATGATCGATCAGGAACGGAAGTTCAATTTCTGTTATTGAG 660
Db |||||
674 TCTGGGAAGGAGTGCCTCATGATCGATCAGGAACGGAAGTTCAATTTCTGTTATTGAG 733
Qy 661 GACCCATCAATCAAGATTTCTGAGCTCAGATACAAATGCTGGCATTAATGCTAGCTTCTG 720
Db |||||
734 GACCCATCAATCAAGATTTCTGAGCTCAGATACAAATGCTGGCATTAATGCTAGCTTCTG 793
Qy 721 CTGCTCAGATAGCTCTCTGAAATTTGAACCAAACTGACCTTTTCAAGATCAATGCGAA 780
Db |||||
794 CTGCTCAGATAGCTCTCTGAAATTTGAACCAAACTGACCTTTTCAAGATCAATGCGAA 853
Qy 781 TGGACCAATCCAAAGCGAATGACCTCTAAATGATGTTGAGGTATTTTCTCAGTTCAAG 840
Db |||||
854 TGGACCAATCCAAAGCGAATGACCTCTAAATGATGTTGAGGTATTTTCTCAGTTCAAG 913

841 ATAGAGAGAAATTAAGAAATGAACAAATTTTGAAGCCTTAGAGCCTTAGGCTGAAGATATCTTT 900
Db |||||
914 ATAGAGAGAAATTAAGAAATGAACAAATTTTGAAGCCTTAGAGCCTTAGGCTGAAGATATCTTT 973
Qy 901 GATGAATCCAAAGCAGATCTCTCTGGGATGCTTCTGGGGGCTGCTGCTGTATATATCAAG 960
Db |||||
974 GATGAATCCAAAGCAGATCTCTCTGGGATGCTTCTGGGGGCTGCTGCTGTATATATCAAG 1033
Qy 961 ATGATGCAAAATCTTACATAGAGTCACTGAGGAGGACCGAGGCTACTGCTGCCACA 1020
Db |||||
1034 ATGATGCAAAATCTTACATAGAGTCACTGAGGAGGACCGAGGCTACTGCTGCCACA 1093
Qy 1021 GGAAGTAATTTGTAAGAACCACTCCCTCAGTCCAGCTGTTTGAAGCTGACACACA 1080
Db |||||
1094 GGAAGTAATTTGTAAGAACCACTCCCTCAGTCCAGCTGTTTGAAGCTGACACACA 1153
Qy 1081 TTCTATTGTTTATCAGAGAGATGATCATCATCTTATTCAGTGGCAAGATTTCTTGGCCT 1140
Db |||||
1154 TTCTATTGTTTATCAGAGAGATGATCATCATCTTATTCAGTGGCAAGATTTCTTGGCCT 1213
Qy 1141 TGA 1143
Db 1214 TGA 1216

RESULT 3
AR082433
LOCUS
DEFINITION
Sequence 30 from patent US 5972886.
AR082433
ACCESSION
AR082433.1
VERSION
GI:10009159
KEYWORDS
Unknown.

ORGANISM Unknown.
REFERENCE 1 (Bases 1 to 1950)
AUTHORS Tsujimoto, M., Iwasa, F., Tsuruoka, N., Nakazato, H., Miura, K.,
Ishida, N., Kurihara, T., Yamaichi, K. and Yamaguchi, N.
TITLE Megakaryocyte differentiation factor
JOURNAL Patent: US 5972886-A 30 26-OCT-1999;
FEATURES Location/Qualifiers
source
1. .1950
BASE COUNT 570 a 384 c 407 g 589 t
ORIGIN
Query Match 100.0%; Score 1143; DB 6; Length 1950;
Best Local Similarity 100.0%; Pred. No. 1e-311;
Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGCCCTCCCTTCTGTCGACGCAATGCGAGATTTTCTTCAACCTGTTTCAGAGAGATGGAT 60
Db 74 ATGGCCCTCCCTTCTGTCGACGCAATGCGAGATTTTCTTCAACCTGTTTCAGAGAGATGGAT 133
Qy 61 GACATCAGGAAATGGAATATGTTCTTTCTCTCTGAGCCTCTTCTGCTGCCCTGCC 120
Db 134 GACATCAGGAAATGGAATATGTTCTTTCTCTCTGAGCCTCTTCTGCTGCCCTGCC 193
Qy 121 CTGCTCCGCTTGGGCGCTCAAGATGATCCCTCTCTCAGATTGATTAAGTTGCTTCATGTT 180
Db 194 CTGCTCCGCTTGGGCGCTCAAGATGATCCCTCTCTCAGATTGATTAAGTTGCTTCATGTT 253
Qy 181 AACACTCCCTCAGGATATGGAACCTCTTCAATAGTCAAGGCTCCAGGCTCCAGCTC 240
Db 254 AACACTCCCTCAGGATATGGAACCTCTTCAATAGTCAAGGCTCCAGGCTCCAGCTC 313
Qy 241 AAAAGAGTTTTTTCTGATATAAATGCATCCACAGGATTAATGATCTCAGCATTTGTAAT 300
Db 314 AAAAGAGTTTTTTCTGATATAAATGCATCCACAGGATTAATGATCTCAGCATTTGTAAT 373
Qy 301 GGGCTTTTCTGTAAGAAAGTGTATGCTTTTATTAAGGACTACATTTGAGTGTGCGAAAAA 360
Db 374 GGGCTTTTCTGTAAGAAAGTGTATGCTTTTATTAAGGACTACATTTGAGTGTGCGAAAAA 433
Qy 361 TTATACGATGCCAAGTGGAGCGAGTGTGCTTTTACGAATCATTTAGAAGACATAGACGT 420
Db 434 TTATACGATGCCAAGTGGAGCGAGTGTGCTTTTACGAATCATTTAGAAGACATAGACGT 493
Qy 421 AATATTAAATAGTGGTGTAAATGAACACATGCGCAAAATCAAGAACCTGATTGTGAA 480
Db 494 AATATTAAATAGTGGTGTAAATGAACACATGCGCAAAATCAAGAACCTGATTGTGAA 553
Qy 481 GGTGGCATTAAGCTCATCTGCTGTAAATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 540
Db 554 GGTGGCATTAAGCTCATCTGCTGTAAATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 613
Qy 541 TGGCAATCAGGCTTCCACAGAGCGAATCAATTAATTTGCCAATTTCAAAATCTCCCAAGTGC 600
Db 614 TGGCAATCAGGCTTCCACAGAGCGAATCAATTAATTTGCCAATTTCAAAATCTCCCAAGTGC 673
Qy 601 TCTGGGAAGGAGTGCCTCATGATCGATCAGGAACGGAAGTTCAATTTCTGTTATTGAG 660
Db 674 TCTGGGAAGGAGTGCCTCATGATCGATCAGGAACGGAAGTTCAATTTCTGTTATTGAG 733
Qy 661 GACCCATCAATCAAGATTTCTGAGCTCAGATACAAATGCTGGCATTAATGCTAGCTTCTG 720
Db 734 GACCCATCAATCAAGATTTCTGAGCTCAGATACAAATGCTGGCATTAATGCTAGCTTCTG 793
Qy 721 CTGCTCAGATAGCTCTCTGAAATTTGAACCAAACTGACCTTTTCAAGATCAATGCGAA 780
Db 794 CTGCTCAGATAGCTCTCTGAAATTTGAACCAAACTGACCTTTTCAAGATCAATGCGAA 853
Qy 781 TGGACCAATCCAAAGCGAATGACCTCTAAATGATGTTGAGGTATTTTCTCAGTTCAAG 840
Db 854 TGGACCAATCCAAAGCGAATGACCTCTAAATGATGTTGAGGTATTTTCTCAGTTCAAG 913

QY 841 ATAGAGAGAAATTATGAAATGAAACAAATTTTGAGAGCCCTAGGCTGAAAGATATCTTT 900
DB 914 ATAGAGAGAAATATGAAATGAAACAAATTTTGAGAGCCCTAGGCTGAAAGATATCTTT 973
QY 901 GATGAATCCAAAGCAGATCTCTCTGGGATTGCTTCGGGGGGTCTCTGTATATATCAAGG 960
DB 974 GATGAATCCAAAGCAGATCTCTCTGGGATTGCTTCGGGGGGTCTCTGTATATATCAAGG 1033
QY 961 ATGATGCACAAATCTTACATAGAGGTCACTGAGAGGGGACCGAGGCTACTGTCGCCACA 1020
DB 1034 ATGATGCACAAATCTTACATAGAGGTCACTGAGAGGGGACCGAGGCTACTGTCGCCACA 1093
QY 1021 GAACTATATTTGTAAGAAACCACTCCCTCAGTCCAGCTGTTAGAGCTGACACCCCA 1080
DB 1094 GAACTATATTTGTAAGAAACCACTCCCTCAGTCCAGCTGTTAGAGCTGACACCCCA 1153
QY 1081 TTCTATTTCTTATCAGGAAGGATGACATCATCTTATTCAGTGGGCAAAATTTCTGGCCT 1140
DB 1154 TTCTATTTCTTATCAGGAAGGATGACATCATCTTATTCAGTGGGCAAAATTTCTGGCCT 1213
QY 1141 TGA 1143
DB 1214 TGA 1216

RESULT 4
E08396 LOCUS 1950 bp RNA linear PAT 29-SEP-1997
DEFINITION cDNA encoding human megakaryocyte growth differentiating factor.
ACCESSION E08396
VERSION E08396.1 GI:2176513
KEYWORDS JP 1994313000-A/20.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Tsujimoto,M., Kurihara,T., Ishida,N., Iwasa,F., Nakazato,H.,
Yamaichi,H., Miura,T., Tsuruoka,N. and Yamaguchi,M.
MEGAKARYOCYTE-PROLIFERATING AND DIFFERENTIATING FACTOR
Patent: JP 1994313000-A 20 NOV-1994;
SUNTORY LTD
OS Homo sapiens (human)
EN JP 1994313000-A/20
PD 08-NOV-1994
PF 16-JUL-1993 JP 199319752
PR 17-JUL-1992 JP 92P 212305, 04-MAR-1993 JP 92P 67339 PI
TSUJIMOTO MASAFUMI, KURIHARA TATSUYA, ISHIDA NOBUHIRO, PI IWASA
FUYUKI
PI NAKAZATO HIROSHI, YAMAICHI HIROZO, MIURA TAKEHISA, PI
TSURUOKA NOBUO,
PI YAMAGUCHI MARE
PC C07K15/14,A61K37/02,C12N5/10,C12N15/19,C12P21/02,C12P21/02,
C12R1:91;
CC strandedness: Double;
CC topology: linear;
CC hypothetical: No;
CC anti-sense: No;
PH Key Location/Qualifiers
FT 1..1950
FT source /organism='Homo sapiens'
FT /cell_line='A431 (human epidermoid FT
FT carcinoma)'
FT 1..73
FT 5'UTR
FT CDS 74..1216
FT /product='Human megakaryocyte growth FT
FT differentiating factor'
FT
FT 3'UTR
FT polyA signal 1217..1950
FT location/Qualifiers 1933..1938.
FT 1..1950

/organism="Homo sapiens"
/mol_type="genomic RNA"
/db_xref="taxon:9606"
BASE COUNT 570 a 384 c 407 g 589 t
ORIGIN
Query Match 100.0%; Score 1143; DB 6; Length 1950;
Best Local Similarity 100.0%; Pred.No.1e-311;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCGCTCCCTTCTGCTGAGCAAAATGAGAGTCTTTCTCTCAACCTGTTTCAGAGAGATGGAT 60
DB 74 ATGCGCTCCCTTCTGCTGAGCAAAATGAGAGTCTTTCTCTCAACCTGTTTCAGAGAGATGGAT 133
QY 61 GACATCAAGGAATGGAATGCTCTTTCTCTCTGAGCCTCTTCGCTCCCTCGCC 120
DB 134 GACATCAAGGAATGGAATGCTCTTTCTCTCTGAGCCTCTTCGCTCCCTCGCC 153
QY 121 CTGCTCCCTTGGGCGCTCAAGATGACTCCCTCTCTCAGATGATTAAGTTGCTTCATGTT 180
DB 194 CTGCTCCCTTGGGCGCTCAAGATGACTCCCTCTCTCAGATGATTAAGTTGCTTCATGTT 253
QY 181 AACACTGCTCAGGATATGGAACCTCTTCTAATAGTCACTCAGGCTCCAGTCTCAACTG 240
DB 254 AACACTGCTCAGGATATGGAACCTCTTCTAATAGTCACTCAGGCTCCAGTCTCAACTG 313
QY 241 AAAGAGTTTTTTCTGATATAATGCAATCCCAAGGATTTATGATCTCAGCATTTCTGAAT 300
DB 314 AAAGAGTTTTTTCTGATATAATGCAATCCCAAGGATTTATGATCTCAGCATTTCTGAAT 373
QY 301 GGGCTTTTTGCTGAAAAGTGTATGGCTTTCTAATAGGACTACATGAGTGTGCGCAAAA 360
DB 374 GGGCTTTTTGCTGAAAAGTGTATGGCTTTCTAATAGGACTACATGAGTGTGCGCAAAA 433
QY 361 TTATACGATGCCAAAGTGAAGGAGTGTACCTTTACGAATCATTTAGAGACACTAGAGCT 420
DB 434 TTATACGATGCCAAAGTGAAGGAGTGTGCTTTAAGAACTATTAGAGACACTAGAGCT 493
QY 421 ATATTAATAGTGGGTTCAAAATGAAACACATGCGCAAAATCAAGAACGTGATGTTGTA 480
DB 494 ATATTAATAGTGGGTTCAAAATGAAACACATGCGCAAAATCAAGAACGTGATGTTGTA 553
QY 481 GTTGCCATAAGCTCATCTCTCTAATGCTGCTGTAATGCTGTGTGTAATGCTGTGTAATG 540
DB 554 GTTGCCATAAGCTCATCTCTCTAATGCTGCTGTAATGCTGTGTGTAATGCTGTGTAATG 613
QY 541 TGGCAATCAGCTTCTCACCAGAGGCGAACCACTAAATGCGCATTTCAATCTCCCAAGTGC 600
DB 614 TGGCAATCAGCTTCTCACCAGAGGCGAACCACTAAATGCGCATTTCAATCTCCCAAGTGC 673
QY 601 TCTGGAGAGGAGTGGCCATGATGATCAGGAACGGAGTTCAATTTGCTGTATTGAG 660
DB 674 TCTGGAGAGGAGTGGCCATGATGATCAGGAACGGAGTTCAATTTGCTGTATTGAG 733
QY 661 GACCCATCAATGAAGATTTCTGAGCTCAGATACATGTTGGGCAATGAACATGACGTTCTG 720
DB 734 GACCCATCAATGAAGATTTCTGAGCTCAGATACATGTTGGGCAATGAACATGACGTTCTG 793
QY 721 CTGCTTGAGATGACCTCTCTGAAATGAAACAAACATGACCTTTCAAGAACTTAATGGA 780
DB 794 CTGCTTGAGATGACCTCTCTGAAATGAAACAAACATGACCTTTCAAGAACTTAATGGA 853
QY 781 TGCAACCAATCCAAAGGCGAATGACCTCTAAGTATGTTGAGGTAATTTTTCCTCAGTTCAAG 840
DB 854 TGCAACCAATCCAAAGGCGAATGACCTCTAAGTATGTTGAGGTAATTTTTCCTCAGTTCAAG 913
QY 841 ATAGAGAGAAATTATGAAATGAAACAAATTTTGAGAGCCCTAGGCTGAAAGATATCTTT 900
DB 914 ATAGAGAGAAATTATGAAATGAAACAAATTTTGAGAGCCCTAGGCTGAAAGATATCTTT 973
QY 901 GATGAATCCAAAGCAGATCTCTCTGGGATTGCTTCGGGGGGTCTCTGTATATATCAAGG 960
DB 974 GATGAATCCAAAGCAGATCTCTCTGGGATTGCTTCGGGGGGTCTCTGTATATATCAAGG 1033


```

QY 1081 TTCCTATTCTTATCAGGAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGGCCT 1140
DB 1154 TTCCTATTCTTATCAGGAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGGCCT 1213
QY 1141 TGA 1143
DB 1214 TGA 1216

RESULT 6
AF027866
LOCUS AF027866 2249 bp mRNA linear PRI 20-OCT-1998
DEFINITION Homo sapiens megin mRNA, complete cds.
ACCESSION AF027866
VERSION AF027866.1 GI:3769372
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2249)
AUTHORS Miyata,T., Nangaku,M., Suzuki,D., Inagi,R., Uragami,K., Sakai,H.,
Okubo,K. and Kurokawa,K.
TITLE A mesangium-predominant gene, megin, is a new serpin upregulated
in IgA nephropathy
J. Clin. Invest. 102 (4), 828-836 (1998)
98376492
MEDLINE 9710452
PUBMED
REFERENCE 2 (bases 1 to 2249)
AUTHORS Toshio,M., Yasuda,Y., Nangaku,M. and Kurokawa,K.
TITLE Direct Submission
JOURNAL Submitted (02-OCT-1997) Medical Science, Tokai University,
Bousaidai, Isehara, Kanagawa 259-11, Japan
REFERENCE 3 (bases 1 to 2249)
AUTHORS Toshio,M., Yasuda,Y., Nangaku,M. and Kurokawa,K.
TITLE Direct Submission
JOURNAL Submitted (19-OCT-1998) Medical Science, Tokai University,
Bousaidai, Isehara, Kanagawa 259-11, Japan
REMARK Sequence update by submitter
COMMENT On Oct 20, 1998 this sequence version replaced gi:3764058.
FEATURES
Location/Qualifiers
1..2249
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_type="mesangial"
365..1507
/codon_start=1
/product="megin"
/protein_id="AAC64506.1"
/db_xref="GI:3769373"
/translation="VASLAANAEEFCNIFRMDNDQNGNVFSSLSLFAALALVEL
GAQDLSQIDKLHVNTPASGYNSNSQSLKEVFSDINASHKDYDLSLVNGL
FAEYGFHDXIECAELKYDAKVERVDFNLEEDREN-NKWNENHGIKNVIGE
GEGSSAVMLVNAVYPKWQSAFTKSETINCHFKPKSGKAVAMHQRKFNLSV
IIDSMLLELYNGVINGVYLLPENDLSEINELTFON-MEWTPRSMHTSKYVEVFF
POFKYENVMKQYLRALGLKIDPDSKADLSIASGRLYISRWVHKSYLIEVTEGT
EATATGSNVIEKQLPQSTLEPRADHPFLVIRKDDILFSGKYSDF"
BASE COUNT 649 a 452 c 461 g 683 t 4 others
ORIGIN
Query Match 100.0%; Score 1143; DB 9; Length 2249;
Best local Similarity 100.0%; Pred. No. 1.1e-311;
Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCTCCCTGCTGAGCAATGACAGATTGGCTTCAACTGTTTCAGAGATGGAT 60
DB 365 ATGGCTCCCTGCTGAGCAATGACAGATTGGCTTCAACTGTTTCAGAGATGGAT 424
QY 61 GACAATCAAGAAATGAAATGTGTCTTTCTCTCTGAGCCCTTGGCTGCCCTGGCC 120
DB 425 GACAATCAAGAAATGAAATGTGTCTTTCTCTCTGAGCCCTTGGCTGCCCTGGCC 484

```

```

QY 121 CTGGTCCGCTTGGCGCTCAGAGTGAATCCCTCTCTCAGATTGATAGTTGCTTCATGTT 180
DB 485 CTGGTCCGCTTGGCGCTCAGAGTGAATCCCTCTCTCAGATTGATAGTTGCTTCATGTT 544
QY 181 AACACTGCCTCAGGATATGGAAACTCTTCTAAATAGTCACTCAGGCTCCAGTCTCAACTG 240
DB 545 AACACTGCCTCAGGATATGGAAACTCTTCTAAATAGTCACTCAGGCTCCAGTCTCAACTG 604
QY 241 AAAAGAGTTTTTCTGATATAATGCAATGCCACAGAGATTATGATCTCAGCAATGTGAAT 300
DB 605 AAAAGAGTTTTTCTGATATAATGCAATGCCACAGAGATTATGATCTCAGCAATGTGAAT 664
QY 301 GGSCCTTTTGTGTAAGAAAGTGTATGGCTTTTATAGAGACTACATTGATGTGCGCAAAA 360
DB 665 GGSCCTTTTGTGTAAGAAAGTGTATGGCTTTTATAGAGACTACATTGATGTGCGCAAAA 724
QY 361 TTATACGATGCCAAAGTCGAGCGAGTTGACTTTTACGAATCATTTAGAGACACTAGAGCT 420
DB 725 TTATACGATGCCAAAGTCGAGCGAGTTGACTTTTACGAATCATTTAGAGACACTAGAGCT 784
QY 421 ATATTAATTAAGTGGCTTGAATAAGAAACACACATGGCAAAATCAAGACGTGATGGTGA 480
DB 785 AATATTAATTAAGTGGCTTGAATAAGAAACACACATGGCAAAATCAAGACGTGATGGTGA 844
QY 481 GTGGCATAGCTCATCTCTGTAATGCTGTGCTGAATGCTGTGTACTTTCAAGGCGCAAG 540
DB 845 GTGGCATAGCTCATCTCTGTAATGCTGTGCTGAATGCTGTGTACTTTCAAGGCGCAAG 904
QY 541 TGGCAATCAGCCTTCAAGAGAGCGAAACCATATAATTTGCCATTTCAAATCTCCAAAGTGC 600
DB 905 TGGCAATCAGCCTTCAAGAGAGCGAAACCATATAATTTGCCATTTCAAATCTCCAAAGTGC 964
QY 501 TCTGGAGAGGCGAGTCCGCCATGATGCATCAGAGCGAGGTTCAATTTTCTGTGTATTCAG 660
DB 965 TCTGGAGAGGCGAGTCCGCCATGATGCATCAGAGCGAGGTTCAATTTTCTGTGTATTCAG 1024
QY 661 GACCCATCAATGAAGATTCTTGAGCTCAGATACATAATGGTGCAATAAACATGTACGTTCTG 720
DB 1025 GACCCATCAATGAAGATTCTTGAGCTCAGATACATAATGGTGCAATAAACATGTACGTTCTG 1084
QY 721 CTGCTTGAGATGACCTCTCTGAAATTTGAAACAAATCTGACCTTTGAGATCTTAATGGAA 780
DB 1085 CTGCTTGAGATGACCTCTCTGAAATTTGAAACAAATCTGACCTTTGAGATCTTAATGGAA 1144
QY 781 TGGACCAATCCAGGCGAATGACCTCTTAAGTATGTTGAGGTATTTTCTCTCAGTTCAG 840
DB 1145 TGGACCAATCCAGGCGAATGACCTCTTAAGTATGTTGAGGTATTTTCTCTCAGTTCAG 1204
QY 841 ATAGAGAGAATTTATGAATGAAACATATTTGAGAGCCCTTAGGCTGAAAGATATCTTT 900
DB 1205 ATAGAGAGAATTTATGAATGAAACATATTTGAGAGCCCTTAGGCTGAAAGATATCTTT 1264
QY 901 GATGAATCCAAAGCAGATCTCTCTGGGATTCCTTCGGGGGGTCTCTGTATATATCAAGG 960
DB 1265 GATGAATCCAAAGCAGATCTCTCTGGGATTCCTTCGGGGGGTCTCTGTATATATCAAGG 1324
QY 961 ATGATGCAAAATCTTACATAGAGTCTCAGGAGGCGCACCGAGGCTTACTGTGCGCAC 1020
DB 1325 ATGATGCAAAATCTTACATAGAGTCTCAGGAGGCGCACCGAGGCTTACTGTGCGCAC 1384
QY 1021 GGAAGTAATATTTAGAAAAAGCAATCCCTCAGTCCACGCTCTTTTAGAGCTGACACCCA 1080
DB 1385 GGAAGTAATATTTAGAAAAAGCAATCCCTCAGTCCACGCTCTTTTAGAGCTGACACCCA 1444
QY 1081 TTCTTATTTGTTATCAGGAAGAGTACATCATCTTATTCAGTGGCAAGTTTCTTGGCCT 1140
DB 1445 TTCTTATTTGTTATCAGGAAGAGTACATCATCTTATTCAGTGGCAAGTTTCTTGGCCT 1504
QY 1141 TGA 1143
DB 1505 TGA 1507

```

RESULT 7

AF105328 1386 bp mRNA linear ROD 19-OCT-2001
 Mus musculus mgsin mRNA, complete cds.
 DEFINITION
 AF105328
 VERSION
 AF105328.1 GI:16269518
 KEYWORDS
 SOURCE
 Mus musculus (house mouse)
 ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1386)
 Nangaku, M., Miyata, T., Suzuki, D., Umezono, T., Hashimoto, T.,
 Wada, T., Yagi, M., Nagano, N., Inagi, R. and Kurokawa, K.

Cloning of rodent mgsin revealed its up-regulation in
 mesangio proliferative nephritis
 Kidney Int. 60 (2), 641-652 (2001)
 21368006
 11473647

REFERENCE
 2 (bases 1 to 1386)
 Miyata, T., Inagi, R. and Kurokawa, K.
 Direct Submission
 Submitted (09-NOV-1999) Institute of Medical Sciences, Tokai
 University, Bohseidai, Isehara, Kanagawa 259-1193, Japan
 Location/Qualifiers

1..1386
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /cell_type="mesangial"
 1..1143
 /notes="human mgsin homolog"
 /codon_start=1
 /product="mgsin"
 /protein_id="AA146768.1"
 /db_xref="GI:16269519"

CDS

/translation="MASLAANAEPGLDFREDSOGNGVFFSLSIPTALTLRL
 GARDCAKIDKALHFNIPSSGNSNPGLOYOLAEVLADINSKDYLSLSTGV
 FAEKVDFHIVIECAENLNAKVERVDNDQDFPKINKWENHGIKKVLGD
 SSUSGAVNLVNAVYFGKWSAFTKIDTUSCRFSPTCPGVNMVHQRRENLT
 IQPPQVLELYHGKISWINEPEDGLCEIESKLSQNLMTNRRKQSVYVFL
 PQPKIKNTYEMHLKIGLKDIFDSSADLGSAGSRLYSKLMHKSFIENSEGT
 BAUAATNNIVEKOLPESTVFRAADRPFLVFKNDIILFTGVKSCP"
 BASE COUNT 407 a 289 c 296 g 393 t 1 others
 ORIGIN

Query Match 68.4%; Score 781.4; DB 10; Length 1386;

Best Local Similarity 80.2%; Pred. No. 1.7e-209;

Matches 917; Conservative 0; Mismatches 226; Indels 0; Gaps 0;

QY 1 ATGGCTCCCTTCTGCACAAATGCAGAGTTTGTCTTCAACCTGTTCAGAGATGGAT 60
 Db |||||
 QY 1 ATGGCTCCCTTCTGCACAAATGCAGAGTTTGTCTTCCCTGAGATCTTCACTGCCCTGACC 60
 Db |||||
 QY 61 GACAATCAAGAAATGGAATGTGTTCTTTCTCTCTGAGCTCTCGCTGCCCTGACC 120
 Db |||||
 QY 61 ACTAGCCAGGAATGGAATGTATTTCTTCTCTCCCTGAGATCTTCACTGCCCTGACC 120
 Db |||||
 QY 121 CTGGTCCCTTGGGCGCTCAAGATGACTCCCTCTCTCAGATGATGAATGTTCTCATGTT 180
 Db |||||
 QY 121 CTAAATCCCTGCTGGGTCTCGAGTGCTGCTGACGTCAGATTGACAGGCACTGCACATT 180
 Db |||||
 QY 181 AACATGCTCCAGATATGGAACCTCTCTTAATAGTCAGTCAGGCTCCAGTCTCAATG 240
 Db |||||
 QY 181 AACATACCATCAAGCAAGAAATCACTCTAATATATCAGCCAGGACTTCAGTATCAATTG 240
 Db |||||
 QY 241 AAAAGAGTTTTTCTGATATAAAGCATCCCAAGGATATGATCTCAGATCTGAT 300
 Db |||||
 QY 241 AAAAGAGTTTTTCTGATATAAAGCATCTCTCATAGGATATGAACTCAGCATGCCACT 300
 Db |||||
 QY 301 GGGCTTTTGTGAAAAGTGTATGGCTTTTATAGGACTACATGAGTGTCCGAAAAA 360
 Db |||||
 QY 301 GGAGTTTTTGCAGAAAAAGTCTATGACTTTTCAATAGAACTACATGAGTGTCTGAAAC 360
 Db |||||

QY 361 TTATACGATGCCAAAGTGGAGCGAGTGTGACTTTACGATCATTTTGAAGACACTAGACCT 420
 Db |||||
 QY 361 TTATACGATGCCAAAGTGGAGCGAGTGTGACTTTACGATCATTTTGAAGATACCAAGATT 420
 Db |||||
 QY 421 AATATTAAATAGTGGTTGAAATGAACAACATCGGCAAAATCAAGACGTTGTTGTTGA 480
 Db |||||
 QY 421 AATATTAAATAGTGGTTGAAATGAACAACATCGGCAAAATCAAGACGTTGTTGTTGA 480
 Db |||||
 QY 481 GGTGGCATTAAGCTCATCTGCTTAATGGTGTGTAATGCTGTGCTTACTTCAAGAGCAAG 540
 Db |||||
 QY 481 AGCAGCTCAGCTCGTGGCTGTCTGCTGTCTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 540
 Db |||||
 QY 541 TGGCATTAAGCTTCACCAAGAGCAAAACCAATAATTTGCAATTTCAAAATCTCCAAAGTGC 600
 Db |||||
 QY 541 TGGAAATCGGCTTCACCAAGACTGATACCTCAGTTGCGGTTTTAGGTTCTCCCAAGTGT 600
 Db |||||
 QY 601 TCTGGGAAGCAGTCGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
 Db |||||
 QY 601 CTTGGAAAGTAGTTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
 Db |||||
 QY 661 GACCATCAATCAAGATTTCTTGAAGTCAAGATCAATGAGTGGCAATAACATGATGATGATG 720
 Db |||||
 QY 661 CACCCACCATGCGAGTTCTTGAAGTCAATGATGATGATGATGATGATGATGATGATGATG 720
 Db |||||
 QY 721 CTGCTGAGATGACCTTCCTGAAATTTGAAATGAAATGAAATGAAATGAAATGAAATGAA 780
 Db |||||
 QY 721 CTGCTGAGGATGGCTTATGAAATTTGAAATGAAATGAAATGAAATGAAATGAAATGAA 780
 Db |||||
 QY 781 TGGACCATCCCAAGCGAGTCACTTAAGTATGTTGAGGATTTTCTCTGATGATGATGATG 840
 Db |||||
 QY 781 TGGACCATGAGGAGGAAATGAAATTTCTGATGATGATGATGATGATGATGATGATGATG 840
 Db |||||
 QY 841 ATAGAGAAATATGAAATGAAATGAAATTTTGAAGCGCTTAGGCTGAAAGATATCTTT 900
 Db |||||
 QY 841 ATAGAGAAATATGAAATGAAATGAAATTTTGAAGCGCTTAGGCTGAAAGATATCTTT 900
 Db |||||
 QY 901 GATGAATCAAGCAGATCTCTCTGGGATGCTTCTGGGGGGTCTGTGTATATATATCAAG 960
 Db |||||
 QY 901 GATGATCAGTGCAGATCTCTCTGGGATGCTTCTGGGGGGTCTGTGTATATATATCAAG 960
 Db |||||
 QY 961 ATGATGCAAAATCTTACATAGAGTCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020
 Db |||||
 QY 961 CTATGCAAGTCAATTCATAGAGTCTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
 Db |||||
 QY 1021 GGAAGTAAATATGTAAGAAAGCAACTCCCTCAGTCCACGCTTTTAGAGTGCACCAACCA 1080
 Db |||||
 QY 1021 GGAAGTAAATATGTAAGAAAGCAACTCCCTCAGTCCACGCTTTTAGAGTGCACCAACCA 1080
 Db |||||
 QY 1081 TTCTATTTTATCAGGAGGAGGATGATATATTTTATCAGTGGCAAGTTTCTTGGCT 1140
 Db |||||
 QY 1081 TTCTATTTTATCAGGAGGAGGATGATATATTTTATCAGTGGCAAGTTTCTTGGCT 1140
 Db |||||
 QY 1141 TGA 1143
 Db |||||
 QY 1141 TGA 1143
 Db |||||

RESULT 8

AF105329 1143 bp mRNA linear ROD 19-OCT-2001
 LOCUS
 DEFINITION
 Rattus norvegicus mgsin mRNA, complete cds.
 AF105329
 ACCESSION
 AF105329.1 GI:16269520
 VERSION
 KEYWORDS
 SOURCE
 Rattus norvegicus (Norway rat)
 ORGANISM
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 1143)
 Nangaku, M., Miyata, T., Umezono, T., Hashimoto, T.,
 Wada, T., Yagi, M., Nagano, N., Inagi, R. and Kurokawa, K.

TITLE	Cloning of rodent messin revealed its up-regulation in mesangio proliferative nephritis
JOURNAL	Kidney Int. 60 (2), 641-652 (2001)
MEOLINE	21368006
PUBMED	11473647
REFERENCE	2. (bases 1 to 1143)
AUTHORS	Miyata T., Nangaku T., Inagi R. and Kurokawa K.
TITLE	Direct Submission
JOURNAL	Submitted (09-NOV-1998) Institute of Medical Sciences, Tokai University, Bohseida, Isehara, Kanagawa 259-1193, Japan
FEATURES	Location/Qualifiers
source	1. .1143
	/organism="Rattus norvegicus"
	/mol_type="mRNA"
	/db_xref="taxon:10116"
	/cell_type="mesangial"
CDS	1. .1143
	/notes="human messin homolog"
	/codon_start=1
	/product="messin"
	/protein_id="AA161769.1"
	/db_xref="GI:16269521"
	/translation="MAGLAAANAETGFLDREMDSQGNVFFSSLSITFALSILRLGARDCARIDKALHFTSPQRQNSNQLQYQLKVLADINSKDYELSTANGFAEVDFHKSVMCAEYNLAKVERVDFNTDQETRFKINKLENTHGKIKKVLGDSLSLSAAMVAVNAYPKGKWSATKTDLSHPSGPGKAVNMHQBEPFLGSLIQEPMQILELOVHGGLISWMLPEDDLSBIESKLQNLMDWNTSKVKKSYNVNVLQPKLEKDMESHKAYGLEDIFVESPADLSIAGSGRLVSKLMHKSLLIVSEBGETATKESNIVSKLIPSTVFIRKNGILIFTGKVSCP"
BASE COUNT	345 a 232 c 261 g 305 t
ORIGIN	
Query Match	66.1%; Score 755.8; DB 10; Length 1143;
Best Local Similarity	78.8%; Pred. No. 2.9e-202;
Matches 901; Conservative	0; Mismatches 242; Indels 0; Gaps 0;
QY	1 ATGGCTCCCTTGTCTGCACAAATGCAGAGTTTTCCTTCAACTGTTGAGAGATGGAT 60
Db	1 ATGGCTCCCTTGTCTGCACAAATGCAGATTTGGCTTCGACTTATTCAGAGATGGAT 60
QY	61 GACATCAAGGAATGGAATGTGTTCTCTCTCTGTAGCTCTTGCCTGCCTGCCGCC 120
Db	61 AGTAGTCAAGAAACGGAATGATTTCTCTCTCCCTGAGCATCTTCACCTGCCCTGAC 120
QY	121 CTGGTCCGCTTGGGCGCTCAAGATGAATCCCTCTCTCAGATGATTAAGTTCATGTT 180
Db	121 CTAAATCCGTTTGGGTGCTCGAGTGACTGTGCACGTCAAGATTCACAGGCCCTGCATTT 180
QY	181 AACATGCTCTCAGATATGGAATCTTCTTAATAGTCAGTCAGGCTCCATCTCAACCTG 240
Db	181 ATCTCCCATCAAGAGGGAAATCAFGAAGCATGAGTCAGTAGGATCGGAATATCAATTT 240
QY	241 AAAAGAGTTTTTCTGATATAATGCATCCCAAGGATTATGATCTCAGCATTTGTGAAT 300
Db	241 AAAAGAGTTCTTGTCTGACATAAATCATCTATCATAGGATTATGAATCAGCATTTGCCAAT 300
QY	301 GGCCTTTTGTGAAAAAGTGTATGGCTTTTCATAGAGCTACATTCAGTGTCCGAAAAA 360
Db	301 GGAGTTTTTGCAGAGAAAGTATTTGATTTTTCATAAGAGCTATATGGAGTGTCTGAAAAA 360
QY	361 TTATACGATGCAAGTGGAGCGAGTTGACTTTACGAATCATTTAGAAGACATACAGCT 420
Db	361 TTATACATGCTAAGTGGAAAGAGCTGATTTTCAATGATATACAGAAACCGAATTT 420
QY	421 AATATTAATAAGTGGGTGAAATGAACAACATGCGAAATTCAGAAAGCTGATTTGGTGA 480
Db	421 AATAATTAATAAGTGTGAAATGAACAACATGCGAAATTCAGAAAGCTGATTTGGGCGAC 480
QY	481 GGTGGCATATAGCTATCTGTATGTGCTGTTGGTGAATGCTGTACTTCAAAAGCGAAG 540
Db	481 AGCAGGCTCAGCTCAACAGCTGTATGTGTGTAGTAAGTGTGTTTACTTCAAAAGCGAAG 540
QY	541 TGGCAATCAGGCTTCCACAGAGCGAAACCATAAATTCGCCATTTCCAAATCTCCCAAGTCG 600

```

BASE COUNT      378 a      243 c      276 g      332 t
ORIGIN
/organism="Rattus norvegicus"
/mol_type="Genomic DNA"
/db_xref="taxon:10116"

```

Query Match	56.1%	Score 755.8	DB 6	Length 1229
Best Local Similarity	79.9%	Pred. No. 2.9e-202		
Matches 901	Conservative 0	Mismatches 242	Indels 0	Gaps 0
QY	1	ATGGCTCCCTTCGTCGACGAATGCGAGATTTCCTTCAACCTGTTGAGAGATCGAT	60	
DB	8	ATGGCTCCCTTCGTCGACGAATGCGAGATTTCCTTCAACCTGTTGAGAGATCGAT	67	
QY	61	GACAAATCAAGGAATGGAAATGTCTTTTCTCTCTGAGCCTCTTCGTCGCCCTGGCC	120	
DB	68	AGTAGTCAAGGAATGGAAATGTCTTCTCTTCTCCCTGAGCATCTTCACTGCCCTGACC	127	
QY	121	CTGGTCCCTTTGGGCTCTGAGATGACTCCCTCTCTGAGATTGATGTTGCTTCATGTT	180	
DB	128	CTAATCCCTTTGGTGTCTCGAGGTGACTGTGCACCTGAGATTGCAAGGCCCTGCATTT	187	
QY	181	AACATGTGCTCAGGATATGCAAACTCTTCTAATAGTCAGTCAGGCTCCAGTCTCAACTG	240	
DB	188	ATCTCCCATCAGACAAGGNAATCATCGACAGTCAGTCAGGCTCCATATCATGTTG	247	
QY	241	AAAGAGTTTTTTCGTATATAAATGCATCCCAAGGATTATGATCTCAGCATTGTGAAT	300	
DB	248	AAAGAGTTTTTTCGTACATAAATCACTCTCATAGGATTATGAATCAGCATTGCCAAT	307	
QY	301	GGCCTTTTTCGTAAGAGTGTATGGCTTTCATAGGACTACATTCAGTGTGCGGAAAA	360	
DB	308	GGAGTTTTTTCGAGAAAGATTGTTGATTTTCATAGAGCTATATGAGGTGTGTGAAGAC	367	
QY	361	TTATACGATGCCAAAGTCGAGCGATTGACTTTACGAATCATTTAAGAGACACTAGAGCT	420	
DB	368	TTATCAATGCTAAGGTGCAAGAGTTTGATTTTCAATGATATACAGAAACCGAGTTT	427	
QY	421	AATATTAATANGTGGTTGAAANTGAAACACATGCGAAATCAAGACGTGATGGTGAA	480	
DB	428	AAAAATTAATAATGGAATTTGAAANTGAAACACATGCGAAATCAAGAGGTGTGGGGAC	487	
QY	481	GGTGCATCAAGCTCATCTCTGTAAATGGTGTGGTGAATGCTGTGATCTTCAAGGCGAG	540	
DB	488	AGCAGCTCAGCTCATCAGCTGTATGGTGTCTAGTGAATGCTGTTTCTTCAAGGCGAG	547	
QY	541	TGGCAATCAGCTTTCACAGAGCGAAACCAATAATTTGCCATTCAAATCTCCCAAGTGC	600	
DB	548	TGGAAATCGCCCTTCCACAGAGTGATACCTTCAGTTGCCATTCAGGTCTCCACGCGT	607	
QY	601	TTCTGGGAAGGCGATCGCCATGATCATCAGGAAGGAGTTCAAATTTGTCTGTATTGAG	660	
DB	608	CTTGGAAAGAGATTATATATGATCATCAAGAACGGAGGTTCAATTTGTCTACCATTCAG	667	
QY	661	GACCATCAATAGAGATTTCTTGAGCTCAGATCAATGGTGGCATAAACATGTACGTTCTG	720	
DB	668	GAGCACCAATGCAGATTTCTTGAGCTACAAATCATGTTGCGATTAAGCATGTACATCATG	727	
QY	721	CTGGCTGAGAAATGACCTCTCTGAAATGGAACAAACATGTACCTTTCAAGATCTATGAAA	780	
DB	728	TTGCCCGAGGATGAOCTATCGCAATTTGAAGACAGCTGAGTTTCCAGATCTAATGAC	787	
QY	781	TGGACCAATCCAGGCGAATGACCTCTTAAGTATGTTCAGGTTATTTTCTCTCAGTTCAAG	840	
DB	788	TGGACCAATAGCAGGAGATGNAATCTCAATATGTGATGTGTCTTCTCCCGAGTTCAAG	847	
QY	841	ATAGAGAGAAATTATGAAATGAACAATATTTGAGAGCCCTAGGCTGAAAGATACTTTT	900	
DB	848	ATAGAGAAAGATTATGAATCAGGAGCCACTGAAATCTGTAGGCTTGGAAAGATCTTT	907	
QY	901	GATGAATCCAAAGCAGATCTCTCTGGATGTCTTCGGGGGGTCTGTGTATATATCAAG	960	
DB	908	GTTGAGTCCAGGCTCATCTCTGTGGAATGCGCTCTGGAGGTGCTCTATGATATCAAG	967	

QY	961	ATGATCGACAAATCTTACATAGAGTCACTGAGGAGGCCACGAGGCTACTGTGGCCACA	1027
Db	968	CTAATTCACAAGTCCCTCATAGAGTCTCAGAGAAGGCCACCGAGCAACTGTGTGCCACA	1027
QY	1021	GGAAGTAAATATGTAGAAAAGCAACTCCCTCAGTCCAGCTGTTTAGAGCTGACCAACCA	1080
Db	1028	GAAAGTAAATATGTAGAAAAGCAACTCCCTCAGTCCAGCTGTTTAGAGCTGACCAACCA	1080
QY	1081	TTCTTATTGTTATCAGGAAGSAGTACATCATCTTATTTCAGTGGCAAAAGTTTCTTGCCTT	1140
Db	1088	TTTCTGTTGTCATAGGAAGSAGTACATCATCTTATTTCAGTGGCAAAAGTTTCTTGCCTT	1147
QY	1141	TGA 1143	
Db	1148	TGA 1150	
RESULT 10			
3D094241			
LOCUS			
DEFINITION	B0094241	Model animals of nephropathy proliferating mesangial cells.	
ACCESSION	B0094241		
VERSION	B0094241.1	GI:22639829	
KEYWORDS	WO 0124628-A/2.		
SOURCE	Rattus norvegicus (Norway rat)		
ORGANISM	Rattus norvegicus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;		
	Rattus.		
REFERENCE	1 (bases 1 to 1229)		
AUTHORS	Miyata,T.		
TITLE	Model animals of nephropathy proliferating mesangial cells		
JOURNAL	Patent: WO 0124628-A 2 12-APR-2001;		
	TOSHIO MIYATA, KIYOSHI KUROKAWA		
COMMENT	OS Rattus norvegicus (rat)		
	EN WO 0124628-A/2		
	PD 12-APR-2001		
	PF 06-OCT-2000 WO 2000JP006988		
	FR 06-OCT-1999 JP 99P 285736		
	PI TOSHIO MIYATA		
	PC A01KG7/027,A61P13/12,A61K45/00,C12N15/12,C12N15/85,G01N33/15,		
	CC G01N33/50		
	CC		
FEATURES	Key	Location/Qualifiers	
source	CDS	(8)..(1147)	
	1..1229		
	/organism="Rattus norvegicus"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:10116"		
BASE COUNT	377 a 242 c 274 g 330 t	6 others	
ORIGIN			
Query Match	65.9%;	Score 753;	DB 6; Length 1229;
Best Local Similarity	78.5%;	Pos. No. 1,8e-201;	
Matches 897;	Conservative	0; Mismatches 246;	Indels 0; Gaps 0;
QY	1	ATGGCTCCCTCTGCTGCGAGAAATGCAAGTATTTGCTTCAACCTGTTACAGAGATGGAT	60
Db	8	ATGGCTCCCTCTGCTGCGAGAAATGCAAGTATTTGCTTCAACCTGTTACAGAGATGGAT	67
QY	61	GACATCAAGAAATGGAAATGTTCTTTTCTCTCTGAGCCTCTTCGTCGCCCTGCC	120
Db	68	AGTAGTCAAGAAATGGAAATGTTCTTTCTCTCTGAGCCTCTTCGTCGCCCTGCC	127
QY	121	CTGGTCCGCTTGGGCGCTCAAGATGACTCCCTCTCTCAGATTGATAAGTTGCTTCATGTT	180
Db	128	CTAAATCCGTTTGGTCTCGAGGTGCTGTNNNGTCAAGTTGACAGGCCCTCGACTTT	187
QY	181	ACATCGCTCAGGATATGGAACCTCTTCTAATAGTCAGTCAGGCTCCAGTCTCACTG	240
Db	188	ATCTCCCATCAAGCAAGGAATTCATCGAACAGTTCAGTAGACTTGCATATCAATG	247

241 AAAAGAGTCTTTCTGATATATAAATGATCCCAAGGATATGATCTCAGCATGTGAAT 300
Db AAAAGAGTCTTTCTGATATATAAATGATCCCAAGGATATGATCTCAGCATGTGAAT 307
QY GGGCTTTTCTGATATATAAATGATCCCAAGGATATGATCTCAGCATGTGAAT 360
Db GGGCTTTTCTGATATATAAATGATCCCAAGGATATGATCTCAGCATGTGAAT 367
QY TTATAGGATCCCAAGGATATGATCTCAGCATGTGAATGATCTCAGCATGTGAAT 420
Db TTATAGGATCCCAAGGATATGATCTCAGCATGTGAATGATCTCAGCATGTGAAT 427
QY AATATATATAGTGGTGGTGAATGATGATGATGATGATGATGATGATGATGATGAT 480
Db AATATATATAGTGGTGGTGAATGATGATGATGATGATGATGATGATGATGATGAT 487
QY GGTGAT 540
Db GGTGAT 547
QY TGGCAATCAGGCTTCCCAAGGATATGATCTCAGCATGTGAATGATCTCAGCATGTGA 600
Db TGGCAATCAGGCTTCCCAAGGATATGATCTCAGCATGTGAATGATCTCAGCATGTGA 607
QY TCTGGGAGGAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db TCTGGGAGGAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 667
QY GAGCCATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Db GAGCCATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 727
QY CTGCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db CTGCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 787
QY TGGACCATCCAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Db TGGACCATCCAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 847
QY ATAGAGAGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
Db ATAGAGAGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 907
QY GATGATCCAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
Db GATGATCCAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 967
QY ATGATGACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
Db ATGATGACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1027
QY GGAAGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
Db GGAAGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1087
QY TTCCTATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
Db TTCCTATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1147
QY TGA 1143
Db TGA 1150

RESULT 11
BD094242
LOCUS BD094242 1147 bp DNA linear PAT 27-AUG-2002
DEFINITION Model animals of nephropathy proliferating mesangial cells.
ACCESSION BD094242
VERSION BD094242.1 GI:22639830
KEYWORDS WO 0124628-A/3.

SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1147)
AUTHORS Miyata, T.
TITLE Model animals of nephropathy proliferating mesangial cells
JOURNAL Patent: WO 0124628-A 3 12-APR-2001;
TOSHIO MIYATA, KIYOSHI KUROKAWA
COMMENT OS Mus musculus (mouse)
PN WO 0124628-A/3
PD 12-APR-2001
PE 06-OCT-2000 WO 2000JP006988
PR 06-OCT-1999 JP 99P 285736
PI TOSHIO MIYATA
PC A01K07/027, A61P13/12, A61K45/00, C12N15/12, C12N15/85, G01N33/15,
PC G01N33/50
CC CC
Key Location/Qualifiers
FT CDS (1)..(1104).
FEATURES
source
1..1147
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
BASE COUNT 351 a 239 c 250 g 307 t
ORIGIN
Query Match 55.5%; Score 748.6; DB 6; Length 1147;
Best Local Similarity 79.8%; Pred. No. 3.2e-200;
Matches 883; Conservative 0; Mismatches 224; Indels 0; Gaps 0;
QY 37 TTCACCTGTTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 96
Db 1 TTCGACTTATTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
QY 97 CTGAGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 156
Db 61 CTGAGCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
QY 157 CAGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 216
Db 121 CAGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
QY 217 CAGTCAGGCTCAGTCCTCAGTCCTCAGTCCTCAGTCCTCAGTCCTCAGTCCTCAGTCCT 276
Db 181 CAGTCAGGCTCAGTCCTCAGTCCTCAGTCCTCAGTCCTCAGTCCTCAGTCCTCAGTCCT 240
QY 277 GATTATGATCTCAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 336
Db 241 GATTATGATCTCAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 337 GACTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 396
Db 301 AACTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
QY 397 AATCATTTTGAACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 456
Db 361 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
QY 457 AAAATCAAGAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 516
Db 421 AAGATCAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 517 AATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 576
Db 481 AAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 577 TGGCATTTTCAATCTCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 636
Db 541 TGGCTGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
QY 637 AAGTTCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 696


```

/clone="RP11-79D21"
/clone_lib="RPC1-11 Human Male BAC"
599. .635
/rpt_family=" (T)n"
repeat_region
1040. .1077
/rpt_family="AT_rich"
repeat_region
1411. .1747
/rpt_family="L1P"
repeat_region
2632. .2992
/rpt_family="MLT1A1"
repeat_region
3115. .3138
/rpt_family="AT_rich"
repeat_region
complement(3487..3828)
/rpt_family="L1MC4a"
repeat_region
3834. .4338
/rpt_family="MLT1C"
repeat_region
complement(4339..5397)
/rpt_family="L1MC4a"
repeat_region
complement(5742..5899)
/rpt_family="MER5B"
repeat_region
6499. .6816
/rpt_family="HALL1"
repeat_region
complement(7416..7621)
/rpt_family="MER3"
repeat_region
complement(8163..9090)
/rpt_family="L2"
repeat_region
9276. .9419
/rpt_family="MIR"
repeat_region
9598. .9658
/rpt_family="MER81"
repeat_region
9659. .10033
/rpt_family="MLT1J"
repeat_region
13986. .14009
/rpt_family="AT_rich"
repeat_region
14011. .14115
/rpt_family="(TA)n"
repeat_region
complement(14942..15044)
/rpt_family="MIR3"
repeat_region
15156. .15298
/rpt_family="L2"
repeat_region
15592. .15773
/rpt_family="L1M4C"
repeat_region
15974. .18310
/rpt_family="L1PA16"
repeat_region
16373. .16374
/notes="1328 bp of bacterial transposon insertion in BAC
excised between these 2 bp"
18311. .18331
/rpt_family="(TAAA)n"
repeat_region
18332. .19566
/rpt_family="L1PA16"
repeat_region
19567. .19587
/rpt_family="(CA)n"
repeat_region
19588. .20042
/rpt_family="L1PA16"
repeat_region
20043. .20064
/rpt_family="AT_rich"
repeat_region
complement(21259..21386)
/rpt_family="MIR3"
repeat_region
22395. .22420
/rpt_family="AT_rich"
repeat_region
complement(22788..22948)
/rpt_family="FRM"
repeat_region
complement(23843..24253)
/rpt_family="WSTA"
repeat_region
complement(24678..25041)
/rpt_family="MLT1A2"
repeat_region
25054. .25094
/rpt_family="HALL1"
repeat_region
25214. .25358
/rpt_family="L1M4C"
repeat_region
complement(25602..25854)
/rpt_family="MIR"

```

```

repeat_region 26746. .26796
/rpt_family="AT_rich"
repeat_region 27451. .27484
/rpt_family="(CAAAA)n"
repeat_region complement(28183..28219)
/rpt_family="MIR"
repeat_region complement(28249..28325)
/rpt_family="MIR"
repeat_region complement(29372..29459)
/rpt_family="MIR"
repeat_region 29539. .30003
/rpt_family="MER66B"
repeat_region 30596. .30624
/rpt_family="AT_rich"
repeat_region 31418. .31829
/rpt_family="L2"
repeat_region 31846. .31915
/rpt_family="(TGAA)n"
repeat_region complement(32019..32103)

Query Match 34.9%; Score 399.4; DB 9; Length 157284;
Best Local Similarity 99.8%; Pred. No. 2.8e-101;
Matches 400; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 743 AAATTGAAACAACTGACCTTCAGAACTCTAATGATGACCAATCCAGGCGAATGA 802
Db 126113 AGATTGAAACAACTGACCTTCAGAACTCTAATGATGACCAATCCAGGCGAATGA 126054

QY 803 CCTCTAAGTATGTTGAGGTATTTTCTCAGTTCAGATAGAGAAGATTTATGAATGA 862
Db 126053 CCTCTAAGTATGTTGAGGTATTTTCTCAGTTCAGATAGAGAAGATTTATGAATGA 125994

QY 863 ALCATATTTGAGAGCCCTAGGCTGAAGATATCTTTGATGATCAAGAGATCTCT 922
Db 125993 AACATATTTGAGAGCCCTAGGCTGAAGATATCTTTGATGATCAAGAGATCTCT 125934

QY 923 CTGGGATTTGTTCTGGGGGGTCTCTGTATATATCAAGATGATGACCAATCTTACATAG 982
Db 125933 CTGGGATTTGTTCTGGGGGGTCTCTGTATATATCAAGATGATGACCAATCTTACATAG 125874

QY 983 AGGTCACTGAGAGGCGACCGAGGCTACTGCTGCCACAGGAAGTATATTTAGAAAGC 1042
Db 125873 AGGTCACTGAGAGGCGACCGAGGCTACTGCTGCCACAGGAAGTATATTTAGAAAGC 125844

QY 1043 AACTCCCTCAGTCCAGCTGTGTTAGAGTGAACACCCATTCCTATTGTTGATCAGGAAGG 1102
Db 125813 AACTCCCTCAGTCCAGCTGTGTTAGAGTGAACACCCATTCCTATTGTTGATCAGGAAGG 125754

QY 1103 ATGACATCATCTTATTCAGTGGCAAGTTTCTTGCCCTTGA 1143
Db 125753 ATGACATCATCTTATTCAGTGGCAAGTTTCTTGCCCTTGA 125713

RESULT 13
AC027524
LOCUS 189092 bp DNA linear HTG 27-MAR-2003
DEFINITION Homo sapiens chromosome 18 clone RP11-317G1 map 18, 3 unordered
pieces.
ACCESSION AC027524
VERSION AC027524.4 GI:21307437
KEYWORDS HTG; HTGS PHASE1; HTGS_FULLTOP; HTGS_CANCELLED.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 189092)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 18, clone RP11-317G1
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 189092)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,I., Boukhgalter,B., Brown,A., Burkett,G.,

```

FEATURES	Location/Qualifiers
source	1..189092
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="18"
	/map="18"
	/clone="RP11-317G1"
	/clone_lib="RP11-11 Human Male BAC"
BASE COUNT	57107 a 35683 c 36719 g 59369 t 214 others
ORIGIN	
Query Match	34.9%; Score 399.4; DB 2; Length 189092;
Best Local Similarity	99.8%; Pred. No. 2.8e-101;
Matches 400; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
QY	743 AAATGAAACAAATGACCTTTCAGATCTTAATGATGGACCAATCAAGGGCAATGA 802
Db	96044 AGATTGAAACAAATGACCTTTCAGATCTTAATGATGGACCAATCAAGGGCAATGA 96103
QY	803 CCTCTAGATATCTTGAGGTATTTTTCTCAGTTCAGATAGAGAAGATTATGAAATGA 862
Db	96104 CCTCTAGATATCTTGAGGTATTTTTCTCAGTTCAGATAGAGAAGATTATGAAATGA 96163
QY	863 AACAAATATTTGAGAGCCCTAGGGCTGAAGCATATCTTTGATGAATCCAAAGCAGATCTCT 922
Db	96164 AACAAATATTTGAGAGCCCTAGGGCTGAAGCATATCTTTGATGAATCCAAAGCAGATCTCT 96223
QY	923 CTGGGATCTCTTCGGGGGGTCTCTGTATATATCAAGGATGATGCACAAATCTTACATAG 982
Db	96224 CTGGGATCTCTTCGGGGGGTCTCTGTATATATCAAGGATGATGCACAAATCTTACATAG 96283
QY	983 AGGTCTACTCAGAGGGGCAACGGGGCTACTCTGTCACAGGAGCTATATTCTAGAAAGC 1042
Db	96284 AGGTCTACTCAGAGGGGCAACGGGGCTACTCTGTCACAGGAGCTATATTCTAGAAAGC 96343
QY	1043 AACTTCCTCTCAGTCACACCTGTTTAGAGTGCACACCCATCTCTATTGTGTATCAGNAGG 1102
Db	96344 AACTTCCTCTCAGTCACACCTGTTTAGAGTGCACACCCATCTCTATTGTGTATCAGNAGG 96403
QY	1103 ATGACATCATCTTATTCAGTGGCAAGTTTCTTTCGCCCTTGA 1143
Db	96404 ATGACATCATCTTATTCAGTGGCAAGTTTCTTTCGCCCTTGA 96444

[illegible]

```

Query Match      26.8%;   Score 304;  DB 6;  Length 316;
Best Local Similarity  95.4%;
Matches 315;  Conservative  0;  Mismatches  1;  Indels  1;  Gaps  1;

822  ATTTTTCCTCAGTTCAAGTAGAGAGAGATTATGAATGAACAACTATTGTGAGGCCCT 881
      |||||
1  ATTTTTCCTCAGTTCAAGTAGAGAGAGATTATGAATGAACAACTATTGTGAGGCCCT 60

```

*	1	85659: contig of 8559 bp in length
*	85660	85759: gap of 100 bp
*	85660	163014: contig of 77355 bp in length
*	163015	163114: gap of 100 bp
*	163115	189092: contig of 25978 bp in length

QY 882 AGGCTGAAGATATCTTTGATGATCCAAAGCAGATCTCTCGGAGTTGCTTCGGGGGG 941
 Db |||||
 QY 61 AGGCTGAAGATATCTTTGATGATCCAAAGCAGATCTCTCGGAGTTGCTTCGGGGGG 120
 Db |||||
 QY 942 TCGTCTGTATATATCAAGGATGATGCACAAATCTTACATAGAGTCACTGAGGAGGCAC 1001
 Db |||||
 QY 121 TCGTCTGTATATCAAGGATGATGCACAAATCTTACATAGAGTCACTGAGGAGGCAC 180
 Db |||||
 QY 1002 CGAGGCTACTGTCGCACAGGAAGTAATATTTGAAAGCAATCCCTCAGTCCACGCT 1061
 Db |||||
 QY 181 CGAGGCTACTGTCGCACAGGAAGTAATATTTGAAAGCAATCCCTCAGTCCACGCT 240
 Db |||||
 QY 1062 GTTTAGAGCTGACCCACCATCTCTATTTGTTATCAGGAGGATGATCATCTATTATCAG 1121
 Db |||||
 QY 241 GTTTAGAGCTGACCCACCATCTCTATTTGTTATCAGGAGG-TCATCATCTATTATCAG 299
 Db |||||
 QY 1122 TGGCAAAAGTTTCTTGCC 1138
 Db |||||
 QY 300 TGGCAAAAGTTTCTTGCC 316
 Db |||||

RESULT 15

AC125314 121141 bp DNA linear ROD 26-MAR-2003
 DEFINITION Mus musculus chromosome 1 clone RP24-361P7, complete sequence.
 AC125314
 ACCESSION
 VERSION AC125314.3 GI:29244780
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Eukaryota; Chordata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 TITLE 1. (bases 1 to 121141)
 JOURNAL McPherson, J.D. and Waterston, R.H.

REFERENCE
 AUTHORS The sequence of Mus musculus clone
 TITLE Unpublished
 JOURNAL 2. (bases 1 to 121141)

REFERENCE
 AUTHORS McPherson, J.D. and Waterston, R.H.
 TITLE Direct Submission
 JOURNAL 3. (bases 1 to 121141)

REFERENCE
 AUTHORS Submitted (22-JUN-2002) Genome Sequencing Center, 4444 Forest Park
 TITLE Parkway, St. Louis, MO 63108, USA
 JOURNAL 4. (bases 1 to 121141)

REFERENCE
 AUTHORS McPherson, J.D. and Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (03-DEC-2002) Genome Sequencing Center, 4444 Forest Park
 TITLE Parkway, St. Louis, MO 63108, USA

REFERENCE
 AUTHORS McPherson, J.D. and Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (26-MAR-2003) Genome Sequencing Center, 4444 Forest Park
 TITLE Parkway, St. Louis, MO 63108, USA

REFERENCE
 AUTHORS On Mar 26, 2003 this sequence version replaced gi:26006669.
 TITLE
 JOURNAL

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

Matches 325; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
 QY 743 AAATTGAAAACAAACTGACCTTTTCAGAAATCTAATGGAATGGACCAATCCAAAGGCGAATGA 802
 Db |||||
 QY 18914 AGATTGAAAAGCAAGCTGAGTTTCCAGAAATCTGATGAGCTGGACCAATAGGAGGAAAATGA 18973
 Db |||||
 QY 803 CCTCTAAGTATGTTGAGGTATTTTTCCTCAGTTCAAGATAGAGGAATTTATGAATGA 862
 Db |||||
 QY 18974 AATCTCAGTATGTGAACGTGTTTCTCCCCCAGTTCAAGATAGAGGAATTTATGAATGA 19033
 Db |||||
 QY 863 AACAAATATTGAGAGCCCTTAGGGCTGAAAGATATCTTTGATGAATCCAAAGCAGATCTCT 922
 Db |||||
 QY 19034 CGACCACTTGAATCTTTAGGCTTGAAGATATCTTTGATGAGTCCAGTSCAGATCTCT 19093
 Db |||||
 QY 923 CTGGGATTCCTTCGGGGGTGCTCTGTATATATATCAAGGATGATGCAAAATCTTTCATAG 982
 Db |||||
 QY 19094 CTGGAATTGCTCTGGAGTCTCTCTACGTATCAAAAGCTAATGCACAGTCAATTCATAG 19153
 Db |||||
 QY 983 AGGTCACTGAGGAGGCGACCGAGGCTACTCTGCCACAGGAGTAATATTGTAGAAAGC 1042
 Db |||||
 QY 19154 AGGTCTCAGAGGAGGCGACTGAAGCCACTGCTGCCACAGAAAATAACATTGTGTGAAGC 19213
 Db |||||
 QY 1043 AACTCCCTCAGTCCACCGCTGTTTAGAGCTCACCCCATTCCTATTTTGTTATCAGGAAG 1102
 Db |||||
 QY 19214 AGCTTCTGAGTCCACAGGTTTCAGAGCGCGACCGCCCTTCTCTGTTGTATCAAGAAGA 19273
 Db |||||
 QY 1103 ATGACATCACTTATTCAGTGGCMAAGTTTCTTGCCCTTGA 1143
 Db |||||
 QY 19274 ATGACATCACTTATTTACTGGCAAGTCTCTTGCTCTGA 19314
 Db |||||

Search completed: December 14, 2003, 02:32:56

Job time : 4381 secs

FEATURES

1. .121141
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="Genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="1"
 /clone="RP24-361P7"

BASE COUNT 35618 a 23200 c 23524 g 37799 t

ORIGIN

Query Match 24.4%; Score 279.4; DS 10; Length 121141;
 Best Local Similarity 81.0%; Pred. No. 2.3e-57;

Result No.	Query	Score			ID	Description
		Match	Length	DB		
1	1143	100.0	1143	20	AAK56712	Human megasin codin
2	1143	100.0	1143	21	AAAR9294	Human megasin nucle
3	1143	100.0	1143	22	AA782438	Human megasin nucle
4	1143	100.0	1950	15	AAO56570	Human megakaryocyt
5	1143	100.0	2249	25	AEK76260	Lung cancer-associ
6	755.8	66.1	1229	22	AAH48181	Rat megasin protein
7	753	65.9	1229	20	AAK56713	Rat megasin DNA, R
8	753	65.9	1229	21	AAK55238	Rat megasin nucleot

WPI; 1999-276983/23.
P-PSDB; AAY08254.

PT Megsin protein expressed specifically in mesangial cells
 XX Claim 4; Page 58-61; 100pp; Japanese.
 XX This invention describes the isolation of novel megin nucleic acid and
 CC proteins from human, rat and mouse tissue. This protein is expressed
 CC specifically in mesangial cells. The products of the invention are
 CC useful for the treatment and diagnosis of diseases involving mesangial
 CC cells, such as IgA nephropathy.
 XX
 SQ Sequence 1143 BP; 336 A; 230 C; 261 G; 316 T; 0 other;
 Query Match 100.08; Score 1143; DB 20; Length 1143;
 Best Local Similarity 100.08; Pred. No. 0;
 Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGGCTTCCTGCTGAGCAATGCGAGATGTTTCTTCAACCTGTCAGAGATGGAT 60
 Db 1 ATGGCTTCCTGCTGAGCAATGCGAGATGTTTCTTCAACCTGTCAGAGATGGAT 60
 QY 61 GACAAATCAAGAAATGGAATGTTTCTTCCCTCTGAGCCTCTTCGCTGCGCTGGCC 120
 Db 61 GACAAATCAAGAAATGGAATGTTTCTTCCCTCTGAGCCTCTTCGCTGCGCTGGCC 120
 QY 121 CTGCTGCGCTGGCGCTCAGATGACTCCCTCTCTCAGATGATGATGTTCTCATGTT 180
 Db 121 CTGCTGCGCTGGCGCTCAGATGACTCCCTCTCTCAGATGATGATGTTCTCATGTT 180
 QY 181 AACACTGCTCAGATATGGAATCTTCTTAATAGTCAGTCAGGCTCCAGTCTCAACTG 240
 Db 181 AACACTGCTCAGATATGGAATCTTCTTAATAGTCAGTCAGGCTCCAGTCTCAACTG 240
 QY 241 AAAAGATTTTTCTGATATGATGATGATGATGATGATGATGATGATGATGATGAT 300
 Db 241 AAAAGATTTTTCTGATATGATGATGATGATGATGATGATGATGATGATGATGAT 300
 QY 301 GGGCTTTTCTGMAAAAGTGTATGGCTTCTAAGAGTACATTCAGTGTGCGCAAAA 360
 Db 301 GGGCTTTTCTGMAAAAGTGTATGGCTTCTAAGAGTACATTCAGTGTGCGCAAAA 360
 QY 361 TTATACATGCCAAAGTGGAGTGTGCTTACGAATCAATTTAGAGACACTAGACGT 420
 Db 361 TTATACATGCCAAAGTGGAGTGTGCTTACGAATCAATTTAGAGACACTAGACGT 420
 QY 421 AATATTTATAGTGGTGTGAAATGAAACACATGCGCAAAATCAAGACGTGATGTTG 480
 Db 421 AATATTTATAGTGGTGTGAAATGAAACACATGCGCAAAATCAAGACGTGATGTTG 480
 QY 481 GTGGCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
 Db 481 GTGGCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
 QY 541 TGGCAATCAGCTTCCAGAGGCAAAACATTAATTTGCCATTTCAATCTCCCAAGTGC 600
 Db 541 TGGCAATCAGCTTCCAGAGGCAAAACATTAATTTGCCATTTCAATCTCCCAAGTGC 600
 QY 601 TCTGGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 Db 601 TCTGGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 QY 661 GACCCATCAATGAAGATTTCTGAGCTCAGATCAATGCTGGCATAAACATGATGTTCTG 720
 Db 661 GACCCATCAATGAAGATTTCTGAGCTCAGATCAATGCTGGCATAAACATGATGTTCTG 720
 QY 721 CTGCTGAGATGACCTCTGAAATTTGAAATCAAACTGACCTTTCAGATCTCAATGAA 780
 Db 721 CTGCTGAGATGACCTCTGAAATTTGAAATCAAACTGACCTTTCAGATCTCAATGAA 780
 QY 781 TGGACCAATCCAGGCGATGACCTCTCAAGTATGTTGAGTATTTTCTCAGTCAAG 840
 Db 781 TGGACCAATCCAGGCGATGACCTCTCAAGTATGTTGAGTATTTTCTCAGTCAAG 840
 QY 841 ATAGAGAGAAATTTATGAAATGAAACAAATATTTGAGAGCCCTGAGGCTGAAAGATATCTTT 900

Db 841 ATAGAGAGAAATTTATGAAATGAAACAAATATTTGAGAGCCCTGAGGCTGAAAGATATCTTT 900
 QY 901 GATGAATCCAAAGCAGATCTCTCTGGGATGCTTCGGGGGCTCTCTCTGATATATCAAG 960
 Db 901 GATGAATCCAAAGCAGATCTCTCTGGGATGCTTCGGGGGCTCTCTCTGATATATCAAG 960
 QY 961 ATGATGACAAATCTTACATAGAGTCTCTGAGGAGGACACAGGCTACTGCTGCCACA 1020
 Db 961 ATGATGACAAATCTTACATAGAGTCTCTGAGGAGGACACAGGCTACTGCTGCCACA 1020
 QY 1021 GGAAGTAAATTTGTAGAAAAGCAACTCCCTCAGTCCAGCTGCTTTAGAGCTGACACCCA 1080
 Db 1021 GGAAGTAAATTTGTAGAAAAGCAACTCCCTCAGTCCAGCTGCTTTAGAGCTGACACCCA 1080
 QY 1081 TTCCTATTTTATCAGGAAGATGACATCATCTTATTCAGTGGGAAAGTTCTTGGCCT 1140
 Db 1081 TTCCTATTTTATCAGGAAGATGACATCATCTTATTCAGTGGGAAAGTTCTTGGCCT 1140
 QY 1141 TGA 1143
 Db 1141 TGA 1143

RESULT 2
 AAA9294
 ID AAA9294 standard; cDNA; 1143 BP.
 XX
 AC AAA9294;
 XX
 DT 30-JAN-2001 (first entry)
 XX
 DE Human megin nucleotide sequence SEQ ID NO:1.
 KW Megsin; mesangium-predominant gene; serpin regulated; nephropathy;
 KW IgA; immunoglobulin A; detection; renal function; renal disorder;
 KW diagnosis; biological sample; blood; urine; ss.
 OS Homo sapiens.
 XX
 PN W0200057189-A1.
 XX
 PD 28-SEP-2000.
 XX
 PF 17-MAR-2000; 2000WO-JP01646.
 XX
 PR 19-MAR-1999; 99JP-0075305.
 PR 28-OCT-1999; 99JP-0306623.
 XX
 PA (KURO/) KUROKAWA K.
 PA (FUSO) FUSO PHARM IND LTD.
 PA (MIYA/) MIYATA T.
 XX
 PI Miyata T;
 DR WPI; 2000-611542/58.
 DR P-PSDB; AAB24142.
 XX
 PT Evaluating renal function comprises assaying megin protein in
 PT biological sample -
 XX
 PS Example 5; Page 62-66; 93pp; Japanese.
 XX
 CC The present invention describes a method for evaluating renal function.
 CC The method comprises assaying megin protein in biological sample. Also
 CC described are: (1) use of an anti-megsin protein antibody for diagnosing
 CC renal function; and (2) a kit for detecting megin protein comprising:
 CC (a) anti-megsin protein antibody attached to solid magnetic particles;
 CC (b) direct or indirect fixing for the antibody to the particles; and
 CC (c) a magnet. The process is useful for evaluating renal function and
 CC diagnosing renal disorders by assaying megin protein in biological
 CC samples (preferably urine or blood). The process is reproducible and
 CC gives accurate results. The present sequence encodes the human megin

CC protein, which is given in the exemplification of the present invention.

XX SQ Sequence 1143 BP; 336 A; 230 C; 261 G; 316 T; 0 other;

Query Match 100.0%; Score 1143; DB 21; Length 1143;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCCCTCCCTGCTGAGCAAAATGAGAGTTTTCCTCAACCTGCTTCAGAGAGTGGAT 60
DB 1 ATGGCCCTCCCTGCTGAGCAAAATGAGAGTTTTCCTCAACCTGCTTCAGAGAGTGGAT 60
QY 61 GACAATCAAGAAATGAAATGTTTCTCTCTGAGCCCTTGGCTGCCGCCC 120
DB 61 GACAATCAAGAAATGAAATGTTTCTCTCTGAGCCCTTGGCTGCCGCCC 120
QY 121 CTGGCCGCTGGCGCTCAGATGACTCCCTCTCTCAGATTGATGATGATGAT 180
DB 121 CTGGCCGCTGGCGCTCAGATGACTCCCTCTCTCAGATTGATGATGATGATGAT 180
QY 181 AACACTGGCTCAGGATATGAAACTCTTCTAATAGTCAGTCAAGGCTCCAGTCTCACTG 240
DB 181 AACACTGGCTCAGGATATGAAACTCTTCTAATAGTCAGTCAAGGCTCCAGTCTCACTG 240
QY 241 AAAAGAGTTTTTCTGATATAAATGATGATGATGATGATGATGATGATGATGAT 300
DB 241 AAAAGAGTTTTTCTGATATAAATGATGATGATGATGATGATGATGATGATGAT 300
QY 301 GGGCTTTTGTGAAAGGTGATGGCTTCTAATAGTCAATGATGATGATGATGATGATGAT 360
DB 301 GGGCTTTTGTGAAAGGTGATGGCTTCTAATAGTCAATGATGATGATGATGATGATGAT 360
QY 361 TTATAGTGCACAAAGTGGAGCGAGTGCATCTTACGAATCATTTAGAGACATGAGGT 420
DB 361 TTATAGTGCACAAAGTGGAGCGAGTGCATCTTACGAATCATTTAGAGACATGAGGT 420
QY 421 AATATTAATAGTGGTGAATAAATAAACAATGAAACATGCAAAATCAAGAACGTGATGGTGA 480
DB 421 AATATTAATAGTGGTGAATAAATAAACAATGAAACATGCAAAATCAAGAACGTGATGGTGA 480
QY 481 GTGGCATAGCTCATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 481 GTGGCATAGCTCATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 541 TGGCAATCAGCTTTCACAGAGCGAAACATTAATGCAATTTCAATCTCCCAAGTGC 600
DB 541 TGGCAATCAGCTTTCACAGAGCGAAACATTAATGCAATTTCAATCTCCCAAGTGC 600
QY 601 TCTGGAGAGGAGTGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 601 TCTGGAGAGGAGTGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 GACCATCAATGAGATCTCTGAGCTCAGTCAATGATGATGATGATGATGATGATGATGAT 720
DB 661 GACCATCAATGAGATCTCTGAGCTCAGTCAATGATGATGATGATGATGATGATGATGAT 720
QY 721 CTGCTGAGATGATCTCTGAAATGAAACAAACTGACCTTCAGATCTAATGAA 780
DB 721 CTGCTGAGATGATCTCTGAAATGAAACAAACTGACCTTCAGATCTAATGAA 780
QY 781 TGGCAATCAAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 781 TGGCAATCAAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 ATAGAGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
DB 841 ATAGAGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
QY 901 GATGAATCAAGCAGATCTCTCTGGGATGCTTCCGGGGTCTCTGATATATCAAG 960
DB 901 GATGAATCAAGCAGATCTCTCTGGGATGCTTCCGGGGTCTCTGATATATCAAG 960
QY 961 ATGATGCAAAATTTACATAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020

DB 961 ATGATGCAAAATTTCTACATAGAGTCTAGAGGGGACCGAGGCTACTGCTGCCACA 1020
QY 1021 GGAAGTAATATTGTAGAAAAGCAACTCCCTCAGTCCACGCTGTTTAGAGCTGACACCCA 1080
DB 1021 GGAAGTAATATTGTAGAAAAGCAACTCCCTCAGTCCACGCTGTTTAGAGCTGACACCCA 1080
QY 1081 TTCTATTTTGTATCAGGAGGATGACATCATCTTATTCAGTGGCAAGTTTCTTGCCT 1140
DB 1081 TTCTATTTTGTATCAGGAGGATGACATCATCTTATTCAGTGGCAAGTTTCTTGCCT 1140
QY 1141 TGA 1143
DB 1141 TGA 1143
RESULT 3
AAP82438 ID AAP82438 standard; cDNA; 1143 BP.
XX AAP82438;
AC AC
XX
DT 10-JUL-2001 (first entry)
XX
DE Human meglin nucleotide sequence.
XX
KW Human; meglin; mesangial cell proliferative nephritis; nephrotropic;
KW transgenic mouse; glomerular disease; animal model; drug screening; ss.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT CDS 1..1143
FT FT /*sg= a
FT FT /product= "meglin"
XX
XX WC200124628-A1.
XX
XX 12-APR-2001.
XX
XX 06-OCT-2000; 2000WO-JP06988.
XX
XX 06-OCT-1999; 99JP-C285736.
XX
XX {KURO/} KUROKAWA K.
XX {MIYA/} MIYAI I.
XX
XX Miyata T;
XX
XX WPI: 2001-300136/31.
XX P-PSDB; AAB83075.
XX
XX Mouse model for mesangial cell proliferative nephritis for development and screening of new treatments -
XX
XX Claim 2; Page 42-44; 62pp; Japanese.
XX
XX The present sequence encodes human meglin. The human meglin coding sequence may be introduced into a mouse to produce an animal model of mesangial cell proliferative nephritis. The symptoms include enlargement of the mesangial base region, sedimentation of an immune complex and an increase in mesangial cells. The animal model is useful for analysing the pathology of chronic glomerular diseases and for screening compositions for prevention and treatment of the diseases.
XX
XX Highly uniform models can be made easily and in large numbers using this method.
XX
XX Sequence 1143 BP; 336 A; 230 C; 261 G; 316 T; 0 other;
Query Match 100.0%; Score 1143; DB 22; Length 1143;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	ATGGCCCTCCCTCTCCAGCAAAATGCAGAGTTTCTCTCAACCTGTTCAGAGAGTGGAT	60
Db	365	ATGGCCCTCCCTCTCCAGCAAAATGCAGAGTTTCTCTCAACCTGTTCAGAGAGTGGAT	424
QY	61	GACAAATCAAGGAATGGAATGTCTTTCTCTCTGAGCCCTCTGCTGCTCCCTGGCC	120
Db	425	GACAAATCAAGGAATGGAATGTCTTTCTCTCTGAGCCCTCTGCTGCTCCCTGGCC	484
QY	121	CTGGTCCTCTGGGCGCTCAAGATGACTCCCTCTCTCAAGATGATAGTGTCTCATGTT	180
Db	485	CTGGTCCTCTGGGCGCTCAAGATGACTCCCTCTCTCAAGATGATAGTGTCTCATGTT	544
QY	181	ACACTGCTCAGGATATGGAATCTCTTAATAGTCAAGTCCAGGCTCCAGTCTCACTG	240
Db	545	ACACTGCTCAGGATATGGAATCTCTTAATAGTCAAGTCCAGGCTCCAGTCTCACTG	604
QY	241	AAAGAGTTTTTCTGATATAATGATCCCAAGGATTAATGATCTCAAGCATTTGAAAT	300
Db	605	AAAGAGTTTTTCTGATATAATGATCCCAAGGATTAATGATCTCAAGCATTTGAAAT	664
QY	301	GGGCTTTTGTGAAAAGTGTATGGCTTTCAATAGGACTACATGAGTGTCCGAAAAA	360
Db	665	GGGCTTTTGTGAAAAGTGTATGGCTTTCAATAGGACTACATGAGTGTCCGAAAAA	724
QY	361	TTATACGATGCGAAAGTGGAGGAGTTGACTTTTACGAATCATTTAGAGCACTAGAGT	420
Db	725	TTATACGATGCGAAAGTGGAGGAGTTGACTTTTACGAATCATTTAGAGCACTAGAGT	784
QY	421	AAATTAATTAAGTGGSTTGAAAATGAACACATGCGAAATCAAGAACGTGTGTTGAA	480
Db	785	AAATTAATTAAGTGGSTTGAAAATGAACACATGCGAAATCAAGAACGTGTGTTGAA	844
QY	481	GGTGCATAGCTCATCTCTCTAATGCTGCTGGTGAATGCTGTACTTCAAGGCAAG	540
Db	845	GGTGCATAGCTCATCTCTCTAATGCTGCTGGTGAATGCTGTACTTCAAGGCAAG	904
QY	541	TGGCAATCAGCTTTCCACAGAGCGAAACCAATATGCAATTTCAAACTCCCAAGTGC	600
Db	905	TGGCAATCAGCTTTCCACAGAGCGAAACCAATATGCAATTTCAAACTCCCAAGTGC	964
QY	601	TCTGGGAAGGCACTGCCATGATGCATCAGGAACGGAAGTTCAATTTGTCTCTATTAG	660
Db	965	TCTGGGAAGGCACTGCCATGATGCATCAGGAACGGAAGTTCAATTTGTCTCTATTAG	1024
QY	661	GACCATCAATGAAGATTTCTGAGTGATGATCAATGTGGCATCAACATGACCTTCTG	720
Db	1025	GACCATCAATGAAGATTTCTGAGTGATGATCAATGTGGCATCAACATGACCTTCTG	1084
QY	721	CTGGCTGAGATGACCTCTCTGAAATGAAACCAAACTGACCTTTCAGAACTTAATGAA	780
Db	1085	CTGGCTGAGATGACCTCTCTGAAATGAAACCAAACTGACCTTTCAGAACTTAATGAA	1144
QY	781	TGGACCAATCCAGGCGAATACCTCTAAGTATGTTGAGTATTTTTTCTCAGTTCAAG	840
Db	1145	TGGACCAATCCAGGCGAATACCTCTAAGTATGTTGAGTATTTTTTCTCAGTTCAAG	1204
QY	841	ATAGAGAGATTTATGAATGAACATATTTGAGAGCCCTAGGCTGGAAGATATCTTT	900
Db	1205	ATAGAGAGATTTATGAATGAACATATTTGAGAGCCCTAGGCTGGAAGATATCTTT	1264
QY	901	GATGAATCCAAAGCAGATCTCTCTGGGATTTGCTTGGGGGTCTCTGTATATATCAAG	960
Db	1265	GATGAATCCAAAGCAGATCTCTCTGGGATTTGCTTGGGGGTCTCTGTATATATCAAG	1324
QY	961	ATGATGCAAAATTTTACATGAGGTCACTGAGGAGGACCCAGGCTACTCTGCCACA	1020
Db	1325	ATGATGCAAAATTTTACATGAGGTCACTGAGGAGGACCCAGGCTACTCTGCCACA	1384
QY	1021	GGAAATATATTGAGAAAGCACTCCCTCAGTCCAGCTGTTAGAGCTGACCAACCA	1080
Db	1385	GGAAATATATTGAGAAAGCACTCCCTCAGTCCAGCTGTTAGAGCTGACCAACCA	1444

RESULT 6
AAH48181
ID AAH48181 standard; DNA; 1229 BP.
XX
AC AAH48181;
XX
DT 21-SEP-2001 (first entry)
XX
Rat megin protein coding sequence.
XX
DE
XX
XX Rat; megin; renal mesangial cell; mesangium proliferative nephritis; ds.
XX
OS Rattus norvegicus.
XX
PN WO200148019-A1.
XX
PD 05-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-JP09251.
XX
PR 28-DEC-1999; 99JP-0373677.
XX
PA (KURO/) KUROKAWA K.
PA (MIYA/) MIYATA I.
XX
PI Miyata T;
XX
DR WPI; 2001-425651/45.
DR P-PSDB; AAG64286.
XX
XX
PT New antibody recognizing a partial sequence of rat megin protein for
diagnosis of mesangium proliferative nephritis -
XX
PS Example 5; Page 52-54; 63pp; Japanese.
XX
CC The present invention relates to a novel antibody which recognises a
peptide consisting of residues 341-354 of rat megin protein. The present
sequence is the coding sequence for rat megin, which was used in the
present invention. Megin is highly expressed in renal mesangial cells
and its level is elevated in mesangium proliferative nephritis. Assay of
the serum or urine level using the antibody is therefore indicative of
this type of disorder.
XX
SQ Sequence 1229 BP; 378 A; 243 C; 276 G; 332 T; 0 other;

Query Match 66.1%; Score 755.8; DB 22; Length 1229;
Best Local Similarity 78.8%; Pred. No. 2.2e-217;
Matches 901; Conservative 0; Mismatches 242; Indels 0; Gaps 0;

QY	1	ATGGCTCCCTTGTGTCAGCAAAATGAGAGTTTGTTCCTCAACCTGTTCAGAGAGTGGAT	60
Db	8	ATGGCTCCCTTGTGTCAGCAAAATGAGAGTTTGTTCCTCAACCTGTTCAGAGAGTGGAT	67
QY	61	GACATCAAGGAATGGAATGTCTTTCTCTCTGAGCCCTCTTCGCTGCTCCCTGGCC	120
Db	68	AGTAGTCAAGGAACCGAAATGTATCTTCTCTCCCTGAGCATCTTCACCTGCCCTGAGC	127
QY	121	CTGGTCCCTTGGCGCTCAAGTGAAGTCTCCCTCTCTCAGATTGATAGTGTCTTCATGTT	180
Db	128	CTAATTCCTTGGTGCTCGAGTGACTGTGACGTCAAGATTGCAAGGCCCTGACCTTT	187
QY	181	AACACTGCTCAGATATGGAACCTCTTCTTAATAGTCAAGTCCAGGCTCCAGTCTCACTG	240

Db 188 ATCTCCCATCAAGACAAGGAATTCATCGAACAGCTAGCTAGGACTGCAATATCAATTG 247
 QY 241 AAAAGAGATTTTCTGATATAAATGATCCACAGAGATATGATCTCAGCATTTGAAT 300
 Db 248 AAAAGAGATTTTCTGATATAAATGATCCACAGAGATATGATCTCAGCATTTGAAT 307
 QY 301 GGGCTTTTCTGATAAAGATGATGCTTTTCAAGAGATATGATCTCAGCATTTGAAT 360
 Db 308 GGAATTTTCTGATAAAGATGATGCTTTTCAAGAGATATGATCTCAGCATTTGAAT 367
 QY 361 TTATACAGTCCAAAGTGGAGGATGATGCTTTTCAAGAGATATGATCTCAGCATTTGAAT 420
 Db 368 TTATACAGTCCAAAGTGGAGGATGATGCTTTTCAAGAGATATGATCTCAGCATTTGAAT 427
 QY 421 AATATTAATAGTGGGTTTGAATTAAGAAACATATGCAAAATCAAGAACTGATGTTGA 480
 Db 428 AAAATTAATAGTGGGTTTGAATTAAGAAACATATGCAAAATCAAGAACTGATGTTGA 487
 QY 481 GGTGGCATAGCTCATCTGCTGTAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
 Db 488 AGCAGCTCAGCTCATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 547
 QY 541 TGGCAATCAGCTCATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
 Db 548 TGGCAATCAGCTCATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 607
 QY 601 TCTGGGAAGGAGTGGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 Db 608 CTTGGGAAGGAGTGGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 667
 QY 661 GACCATCAATCAGCTCATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 Db 668 GACCATCAATCAGCTCATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 727
 QY 721 CTGCTGAGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 Db 728 TTGCTCGAGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 787
 QY 781 TGAACCAATCCAAAGGAGTGGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 Db 788 TGAACCAATCCAAAGGAGTGGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 847
 QY 841 ATAGAGAGATTTATGAATGAACCAATTTTGAAGGAGTGGCCTGCTGCTGCTGCTGCTGCT 900
 Db 848 ATAGAGAGATTTATGAATGAACCAATTTTGAAGGAGTGGCCTGCTGCTGCTGCTGCTGCT 907
 QY 901 GATGAATCCAAAGGAGTGGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 Db 908 GTTGATCCAGGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 967
 QY 961 ATGATGCAAAATCTTACATAGAGTCACTGAGAGGAGGCAAGGAGTGGCCTGCTGCTGCTGCT 1020
 Db 968 CTAATGCAAAATCTTACATAGAGTCACTGAGAGGAGGCAAGGAGTGGCCTGCTGCTGCTGCT 1027
 QY 1021 GGAAGTAAATTTGTAAGAGGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
 Db 1028 GGAAGTAAATTTGTAAGAGGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1087
 QY 1081 TTCTTAATTTGTAAGAGGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
 Db 1088 TTCTTAATTTGTAAGAGGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1147
 QY 1141 TGA 1143
 Db 1148 TGA 1150

RESULT 7

AA56713

ID AA56713 standard; DNA; 1229 BP.

XX

AC AA56713;

XX

DT 14-JUL-1999 (first entry)
 XX
 DE Rat megin DNA.
 XX
 KW Megsin; mesangial cell; treatment; diagnosis; disease; IgA nephropathy;
 human; rat; murine; ss.
 XX
 OS Rattus rattus.
 XX
 FH Key Location/Qualifiers
 FT CDS 8..1150
 FT /+tag= a
 FT /product= "megsin"
 XX
 PN WO9915652-A1.
 PD 01-APR-1999.
 XX
 PF 22-SEP-1998; 58WO-JP04269.
 XX
 PR 22-SEP-1997; 97JP-0275302.
 XX
 PA (KURO/) KUROKAWA K.
 PA (MIYA/) MIYATA T.
 XX
 PT Miyata T;
 XX
 DR WPI; 1999-276983/23.
 DR P-PSDB; AAY08255.
 XX
 XX Megsin protein expressed specifically in mesangial cells
 Claim 4; Page 65-69; 100pp; Japanese.
 XX
 CC This invention describes the isolation of novel megin nucleic acid and
 CC proteins from human, rat and mouse tissue. This protein is expressed
 CC specifically in mesangial cells. The products of the invention are
 CC useful for the treatment and diagnosis of diseases involving mesangial
 CC cells, such as IgA nephropathy.
 XX
 SQ Sequence 1229 BP; 377 A; 242 C; 274 G; 330 T; 6 other;
 Query Match 65.9%; Score 753; DB 20; Length 1229;
 Best Local Similarity 78.5%; Pred. No. 1.5e-216;
 Matches 897; Conservative 0; Mismatches 246; Indels 0; Gaps 0;
 QY 1 ATGGCTCCCTTGGTGGCAGCAATGAGAGTTTGGCTTCAACTGTTTCAGAGAGATGGAT 60
 Db 8 ATGGCTCCCTTGGTGGCAGCAATGAGAGTTTGGCTTCAACTGTTTCAGAGAGATGGAT 67
 QY 61 GACAAATCAGGAAATGGAATGTTGTTTCTCTCTGAGCTTCTTCGTCGCTGGCC 120
 Db 68 AGTAGTCAAGGAAACGGAATGTTTCTCTCTCTGAGCTTCTTCGTCGCTGGCC 127
 QY 121 CTGGTCCGCTTGGGCGCTCAAGATGACTCCCTCTCTCAGATTCATAGTTGCTTCATGTT 180
 Db 128 CTATCCGTTTGGGTGCTCGAGTGTGNNNCGTCAATTCAGAGGCTGCACTTT 187
 QY 181 AACATGCTCAGGATATGAAATCTTCTAATAGTCACTGAGGCTCCAGTCTCAACTG 240
 Db 188 ATCTCCCATCAAGACAAGGAAATTCATCGACAGTCACTAGGACTGCAATATCAATTG 247
 QY 241 AAAAGAGTCTTCTGATATTAATGCAATCCCAAGGATATGATCTCAGCATTTGAAAT 300
 Db 248 AAAAGAGTCTTCTGATATTAATGCAATCCCAAGGATATGATCTCAGCATTTGAAAT 307
 QY 301 GGGCTTTTGTGAAAAAGTGTATGGCTTTTCAATAGGATATGATCTCAGCATTTGAAAT 360
 Db 308 GGGCTTTTGTGAAAAAGTGTATGGCTTTTCAATAGGATATGATCTCAGCATTTGAAAT 367
 QY 361 TTATACAGTCCAAAGTGGAGGATGATGCTTTTCAAGAGATATGATCTCAGCATTTGAAAT 420
 Db 368 TTATACAGTCCAAAGTGGAGGATGATGCTTTTCAAGAGATATGATCTCAGCATTTGAAAT 427

QY 421 AATATTAATAGTGGCTTGAATAATGAACACATGGCAAAATCAAGAACGTGATGGTGAA 480
 Db 428 AAAATTAATTAATGGATTGAAATGAACACATGGCAAAATCAAGAACGTGATGGTGAA 487
 QY 481 GGTGGCATTAAGCTCATCTGCTGTAAATGGTGTGGTGAATGCTGTGACTTCAAGGCAAG 540
 Db 488 ACCAGCTCAGCTCATCAGTGTGATGCTGTGAGTGAATGCTGTGACTTCAAGGCAAG 547
 QY 541 TGGCAATCAGCTTCCAGAGCGAAACCAATTAATGCGATTTCAAAATCTCCCAAGTGC 600
 Db 548 TGGAAATCGCCCTTCCAGAGAGTATACCTCAGTGTGCAATTTCAAGGCTCCCAAGGCT 607
 QY 601 TCTGGGAAGCGCTGCCCATGATCATCAGAACGAAAGTTCAAATTTGCTGTATTCAG 650
 Db 608 CTTGGAAAGCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 657
 QY 661 GACCATCAATGAAGATCTTGGCTCAGATCAATGATGATGATGATGATGATGATGATGATG 720
 Db 668 GAGCCACCAATGCAGATCTTGGCTCAGATCAATGATGATGATGATGATGATGATGATGATG 727
 QY 721 CTGCTGAGATGATCTCTGAAATGAAACAACTGACCTTTGATGATGATGATGATGATGATG 780
 Db 728 ATGCGGAGGATGACCTATCCGAAATGAAAGCAAGCTGAGTTTCCGAAATCTAATGAGC 787
 QY 781 TGGCAATCCAGGCGAATGACCTCTAAGTATGTTGAGTATTTTCTCAGTCAAG 840
 Db 788 TGACCAATAGCAGGAGATGAAATCTCAGTATGATGATGATGATGATGATGATGATGATG 847
 QY 841 ATAGAGAAATATGAATGAACAAATTTGAGAGCCCTAGGCTGGAAGATATCTTT 900
 Db 848 ATAGAGAAATATGAATGAGGAGCCACTTGAATCTGTAGGCTTGGAGACATCTTT 907
 QY 901 GATGATCCAAAGCAATCTCTGGGATGCTTGGGGGTCTGCTGTATATATCAAG 950
 Db 908 GTTGAGTCCAGGCTGATCTGTGGAATGCTCTGAGGCTGAGTCTGTATATATCAAG 957
 QY 961 ATGATGCACAAATCTTACATAGAGTCACTGAGAGGCGACCGAGGCTACTCTCCACA 1020
 Db 968 CTAATGCACAAATCTTACATAGAGTCTGAGAGGCTGAGAGGCGACCGAGGCTACTCTCCACA 1027
 QY 1021 GGAATTAATTTGAAGAACCACTCCCTCAGTCCACGCTCTTTAGAGTCAACCA 1080
 Db 1028 GAAATTAATTTGAAGAACCACTCCCTCAGTCCACGCTCTTTAGAGTCAACGCTCC 1087
 QY 1081 TTCCATTTGATCAGGAGGATGATCATCTTATTCAGTGGCAATTTCTGCT 1140
 Db 1088 TTTCTTTGCTATGAGGAATGGCATCTTATTTACTGGCAAGTCTGCTGCT 1147
 QY 1141 TGA 1143
 Db 1148 TGA 1150

RESULT 8

AAC55238
 ID AAC55238 standard; DNA; 1229 BP.

AC AAC55238;

XX 30-JAN-2001 (first entry)

DE Rat megin nucleotide sequence SEQ ID NO:18.

KW Meglin; mesangium-predominant gene; serpin regulated; nephropathy;
 KW IgA; immunoglobulin A; detection; renal function; renal disorder;
 KW diagnosis; biological sample; blood; urine; ss.

OS Rattus norvegicus.

XX WO200057189-A1.

PN 28-SEP-2000.

Db

XX 17-MAR-2000; 2000WO-JP01646.
 XX 19-MAR-1999; 99JP-0075305.
 PR 28-OCT-1999; 99JP-0306623.
 XX (KURO/) KUROKAWA K.
 PA (FUSO) FUSO PHARM. IND. LTD.
 PA (MIYA) MIYATA T.
 XX Miyata T;
 XX WPI: 2000-611642/58.
 DR P-PSDB; AAB24150.
 XX Evaluating renal function comprises assaying megin protein in
 PT biological sample -
 PT
 PS Disclosure; Page 77-81, 93pp; Japanese.
 XX The present invention describes a method for evaluating renal function.
 CC The method comprises assaying megin protein in biological sample. Also
 CC described are: (1) use of an anti-meglin protein antibody for diagnosing
 CC renal function; and (2) a kit for detecting megin protein comprising:
 CC (a) anti-meglin protein antibody attached to solid magnetic particles;
 CC (b) direct or indirect fixing for the antibody to the particles; and
 CC (c) a magnet. The process is useful for evaluating renal function and
 CC diagnosing renal disorders by assaying megin protein in biological
 CC samples (preferably urine or blood). The process is reproducible and
 CC gives accurate results. The present sequence encodes the rat megin
 CC protein, which is given in the exemplification of the present invention.
 XX
 SQ Sequence 1229 BP; 377 A; 242 C; 274 G; 330 T; 6 other;

Query Match 65.9%; Score 753; DB 21; Length 1229;

Best Local Similarity 78.5%; Pred. No. 1.5e-216;

Matches 897; Conservative 0; Mismatches 246; Indels 0; Gaps 0;

QY 1 ATGGCTCCCTTGTCTGCAACAAATGCGAGTGTGTTGCTTCACTGTTTCAGAGATGAT 60
 Db 8 ATGGCTCCCTTGTCTGCAACAAATGCGAGTGTGTTGCTTCACTGTTTCAGAGATGAT 67
 QY 61 GACATCAAGAAATGGAATGTGTTCTTCTCTGAGCTCTTCTGCTGCTGCTGCTGCT 120
 Db 68 ACTAGTCAAGAAACGGAATCTATTTCTTCTCTGAGCTCTTCTGCTGCTGCTGCTGCT 127
 QY 121 CTGGTCCCTTGGGCGCTCAAGTACGCTCCCTCTCTCAGATTGATAAGTTGCTTCATGTT 180
 Db 128 CTAATCCGTTTGGTGTCTGAGTGTGCTGTTNNCGTCAGATTGACAAGGCTGCACTTT 187
 QY 181 AACCTGCTCAGGATATGGAATCTTCTTAATAGTCACTGAGGCTCCAGTCTCACTG 240
 Db 188 ATCTCCCATCAGACAAAGGAATTCATGCAAGCTGCTAGGACTGCAATATCAATG 247
 QY 241 AAAAGAGTTTCTTCTGATATAAATGCAATCCCAAGGATTTATGATCTCAGATTTGTAAT 300
 Db 248 AAAAGAGTTTCTTCTGATATAAATGCAATCCCAAGGATTTATGATCTCAGATTTGTAAT 307
 QY 301 GGCTTTTCTGTAAGAAAGTGTGCTTTCTAAGAGCTACATGATGATGATGATGATGATG 360
 Db 308 GGAGTTTCTGCAAGAAAGTGTGCTTTGATTTTCTAAGAGCTATATGAGGATGCTGCAAAAC 367
 QY 361 TTATAGATGCAAGAGGAGGAGTGTGCTTTTACGAATCTTTTAGAAGACATAGAGCT 420
 Db 368 TATACAAATGCTAAAGTGAAGAGTGTGATTTTACAAATGATATACAAAGAACCGATTT 427
 QY 421 AATATTAAATAGTGGTGTGAAATGAACACATGCGAAATCAAGAAACGCTGATGTTGTA 480
 Db 428 AATATTAAATAGTGTGAAATGAACACATGCGAAATCAAGAAAGTGTGTTGGGAC 487
 QY 481 GTGGCATAGCTCATCTCTTAATGCTGCTGGTGAATGCTGTGATCTTCAAGGCAAG 540
 Db 488 AGCAGCTCAGCTCATCAGTGTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 547

QY 541 TGGCAATCAGCCTTCACCAAGCAGCAACCAATTAATGGCAATTCABATCTCCCAAGTGC 600
 PA |||||
 DB 548 TGGAAATCGGCTTCACCAAGAGTGAATACCTCTGATTCGCAATTCAGTCTCCCAAGCGT 607
 XX |||||
 QY 601 TCTGGGAAGCGAGTGGCAATGATGCAATCAGGAACGGAAGTTCAATTTCTGTATTGAG 660
 PA |||||
 DB 608 CTTGGAAGAGCAGTAAATATGATGATCAATCAAGACGGAAGTTCAATTTCTGTATTGAG 667
 XX |||||
 QY 661 GACCAATCAATGAGATCTTGAGCTCAGATACAAATGGTGGCAATTAATGATGAGTCTG 720
 PA |||||
 DB 668 GACCAATCAATGAGATCTTGAGCTCAGATACAAATGGTGGCAATTAATGATGAGTCTG 727
 XX |||||
 QY 721 CTGCTGAGATGACCTCTCTGAAATTCGAAATCAAACTGACCTTTTCAATCTAATGGA 780
 PA |||||
 DB 728 TTGCCCGAGATGACCTATCCGAATGGAAGCAGCTGAGTTCCAGATCTAATGGAC 787
 XX |||||
 QY 781 TGGACCAATCCAGGCAATGACCTCTAAGTATGTTGAGTATTTTCTCTCAGTTCAAG 840
 PA |||||
 DB 788 TGGACCAATGAGGCAATGACCTCTAAGTATGTTGAGTATTTTCTCTCAGTTCAAG 847
 XX |||||
 QY 841 ATGAGAGAAATTAATGAATGAACCAATATTTGAGACCTTAGGCTGAAAGATATCTTT 900
 PA |||||
 DB 848 ATGAGAGAAATTAATGAATGAACCAATATTTGAGACCTTAGGCTGAAAGATATCTTT 907
 XX |||||
 QY 901 GATGAATCCAAAGCAGATCTCTCTGGGATTCCTTCGGGGGCTGCTGTATATATCAAG 960
 PA |||||
 DB 908 GTTGATCCAGGCTGATCTGCTGGAATTCCTCTGGAGTCTCTATGATCAAG 967
 XX |||||
 QY 961 ATGATGCAAAATCTTACATAGAGTCTCTGAGAGGCAACGAGGCTACTGTGCCACA 1020
 PA |||||
 DB 968 CTAATGCAAAATCTTACATAGAGTCTCTGAGAGGCAACGAGGCTACTGTGCCACA 1027
 XX |||||
 QY 1021 GGAAGTAAATTTAGAAAGCAATCTCCCTCAGTCCACGCTTTAGAGTCAACACCA 1080
 PA |||||
 DB 1028 GAAAGTAAATTTAGAAAGCAATCTCCCTCAGTCCACGCTTTAGAGTCAACACCA 1087
 XX |||||
 QY 1081 TTCTATTTTATCAGGAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGGCCT 1140
 PA |||||
 DB 1088 TTCTATTTTATCAGGAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGGCCT 1147
 XX |||||
 QY 1141 TGA 1143
 PA |||||
 DB 1148 TGA 1150

RESULT 9

ID AAF82439 standard; cDNA; 1229 BP.
 AC AAF82439;
 XX 10-JUL-2001 (first entry)
 DT
 DE Rat meglin nucleotide sequence.
 XX
 XX Rat; meglin; mesangial cell proliferative nephritis; nephrotropic;
 KW transgenic mouse; glomerular disease; animal model; drug screening; ss.
 XX
 OS Rattus norvegicus.
 XX
 XX Key Location/Qualifiers
 FH 8..1150
 CDS /*tag= a
 FT /product= "meglin"
 ET
 XX W0200124628-AL.
 PN
 XX 12-APR-2001.
 PD
 XX 06-OCT-2000; 2000WO-JP6988.
 PF
 XX 06-OCT-1999; 99JP-0285736.
 PR

XX (KURO/) KUROKAWA K.
 PA (MIYA/) MIYATA T.
 XX Miyata T;
 XX WPI; 2001-300136/31.
 DR P-FSDB; AAB83076.
 XX
 XX Mouse model for mesangial cell proliferative nephritis for development
 PT and screening of new treatments
 XX
 PS Disclosure; Page 46-48; 62pp; Japanese.
 XX
 XX The present sequence encodes rat meglin. The human meglin coding
 CC sequence may be introduced into a mouse to produce an animal model of
 CC mesangial cell proliferative nephritis. The symptoms include
 CC enlargement of the mesangial base region, sedimentation of an immune
 CC complex and an increase in mesangial cells. The animal model is useful
 CC for analysing the pathology of chronic glomerular diseases and for
 CC screening compositions for prevention and treatment of the diseases.
 CC Highly uniform models can be made easily and in large numbers using
 CC this method.
 XX
 SQ Sequence 1229 BP; 377 A; 242 C; 274 G; 330 T; 6 other;
 Query Match 65.9%; Score 753; DB 22; Length 1229;
 Best Local Similarity 78.5%; Pred. No. 1.5e-216;
 Matches 897; Conservative 0; Mismatches 246; Indels 0; Gaps 0;
 QY 1 ATGCGCTCCCTTGTGTCGAGCAAAATGCAAGATTTTGGCTTCACTGTCAGAGATGGAT 60
 DB 8 ATGCGCTCCCTTGTGTCGAGCAAAATGCAAGATTTTGGCTTCACTGTCAGAGATGGAT 67
 QY 61 GACAATCAAGCAAAATGGAATGTGTTTCTCTCTGAGCTCTTCGCTGCTGGCC 120
 DB 68 AATGATCAAGCAAAATGGAATGTGTTTCTCTCTGAGCTCTTCGCTGCTGGCC 127
 QY 121 CTGCTCCGCTTGGCGCTCAAGATGATCCCTCTCTCAGATGATAAGTTGCTTCATGTT 180
 DB 128 CTAATCCGTTTGGTGTCTGAGTGACTGTNNCGTCAGATTGACAAGGCCCTGCATTT 187
 QY 181 AACATGCTCTCAGGATATGGAATCTTCTTAATAGTCACTGAGGCTCAGCTCAACTG 240
 DB 188 ATCTCCCATCAAGCAAGGAATTCATCGAACAGTCACTAGGACTGCAATATCAATTG 247
 QY 241 AAAAGAGTTTTTCTGATATAATGATCATCCACAGGATTTGATCTCAGATTTGAT 300
 DB 248 AAAAGAGTTTTTCTGATATAATGATCATCTCATAGGATNNAACTCAGATTTGCAAT 307
 QY 301 GGCCTTTTGTGAAAAGTATGGCTTTTCTATAAGGACTACATTTAGTGTGCCGAAAA 360
 DB 308 GGAAGTTTTTGCAGAGAAAGTATTTGATTTTCTATAAGGCTATATGAGTGTGCTGAAAAC 367
 QY 361 TTATAGATGTCGCAAGTGGAGCGTGTGACTTTTACCAATCTATTAGAGACATAGAGT 420
 DB 368 TTATAGATGTCGCAAGTGGAGCGTGTGACTTTTACCAATCTATTAGAGACATAGAGT 427
 QY 421 AATATTAATAGTGGTTGAAAATGAAACACATGGCAAAATCAAGACGTTGTTGGTGA 480
 DB 428 AAAATTAATAGTGGTTGAAAATGAAACACATGGCAAAATCAAGAGGTTGTTGGGGAC 487
 QY 481 GTTGGCATAGCTCATCTCTGTAATGCTGCTGTAATGCTGTGTAATGCTGTACTTCAAAGGCAAG 540
 DB 488 AGCAGCTCAGCTCATCAGCTGTCTGATGCTGCTAGTGAATGCTGTTTACTTCAAAGGCAAG 547
 QY 541 TGGCAATCAGCCTTCACCAAGAGGCAACCAATTAATTTGCCATTTCAATCTCCCAAGTGC 600
 DB 548 TGGCAATCAGCCTTCACCAAGAGGCAACCAATTAATTTGCCATTTCAATCTCCCAAGTGC 607
 QY 601 TCTGGGAAGGCAAGTTCGCCCATGATGATCAAGAAAGGAGTTCAATTTGCTGTATTGAG 660
 DB 608 CTTGGGAAGGCAAGTTCGCCCATGATGATCAAGAAAGGAGTTCAATTTGCTGTATTGAG 667

```

QY 661 GACCCATCAATGAAGATCTTACGCTCAGATACAAATGTCGCATAAAACATGATGCTTCG 720
D 668 GAGCCACCAATGACAGATCTTACGCTACAAATATCATGTCGCATGAAGCATGTACATCATG 727
QY 721 CTGCTCAGAGATGACCTCTCTGAAATTTGAACCAACATGACCTTTGAGAACTAATGGA 780
D 728 TTGCCCCGAGATGACCTATCCGAAATTTGAACCAACATGACCTTTGAGAACTAATGGA 787
QY 781 TGGACCAATCCAAAGCGAATGACCTCTAAGTATGTTGAGGATTTTCTCAGTCAAG 840
D 788 TGGCAATAGCAGGAATGAAATCTCAGTATGTAATGTTTCTCCCGAGTCAAG 847
QY 841 ATAGAGAAATTTGAAATGAAGAATTTGAGAGCCCTAGGCTGAAAGATATCTTT 900
D 848 ATAGAGAAATTTGAAATGAGGAGCACTTGAATCTGTAGGCTTGAAGACATCTTT 907
QY 901 GATGAATCCAAAGCAGATCTCTCTGGGATGCTTGGGGGTGCTGTGTATATCAAG 960
D 908 GTTGAATCCAAAGCAGATCTCTCTGGGATGCTTGGGGGTGCTGTGTATATCAAG 967
QY 961 ATGATGACAAATTTTACATAGAGTCACTGAGGAGGCAACCGAGGCTACTGTCGACA 1020
D 968 CTATGACCAAGTCCCTCATAGAGTCTCAGAGAGGCAACCGAGGCTACTGTCGACA 1027
QY 1021 GGAAGTAAATTTGAGAAAGCAATCTCCCTCAGTCAAGCTGCTGTTAGCTGACGACCA 1080
D 1028 GAAAGTAAATTTGAGAAAGCAATCTCCCTCAGTCAAGCTGCTGTTAGCTGACGACCA 1087
QY 1081 TTCCATTTTGTATAGGAGGATGACATCATCTTATTCAGTGGCAAGTTTCTTGCCCT 1140
D 1088 TTCTGTTTGTATAGGAGGATGACATCATCTTATTCAGTGGCAAGTTTCTTGCCCT 1147
QY 1141 TGA 1143
D 1148 TGA 1150

RESULT 10
AA56714
ID AA56714 standard; DNA; 1147 BP.
XX
AC AA56714;
XX
DT 14-JUL-1999 (first entry)
XX
DE Mouse megsin DNA.
XX
KW Megsin; mesangial cell; treatment; diagnosis; disease; IgA nephropathy;
KW human; rat; murine; ss.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
CDS 1..1107
ET /*tag= a
ET /product= "megsin"
FT /note= "Partial sequence, no start codon given"
XX
FN W09915652-A1.
XX
PD 01-APR-1999.
XX
PF 22-SEP-1998; 98WC-JP04269.
XX
PR 22-SEP-1997; 97JP-0275302.
XX
PA (KURO/I) KUROKAWA K.
PA (MIYA/I) MIYATA T.
XX
PI Miyata T;
XX
W21; 1999-276983/23.

```

```

DR P-PSDB; AA509256.
XX
PT Megsin protein expressed specifically in mesangial cells
XX
PS Claim 4; Page 72-76; 100pp; Japanese.
XX
CC This invention describes the isolation of novel megsin nucleic acid and
CC proteins from human, rat and mouse tissue. This protein is expressed
CC specifically in mesangial cells. The products of the invention are
CC useful for the treatment and diagnosis of diseases involving mesangial
CC cells, such as IgA nephropathy.
XX
SQ Sequence 1147 BP; 351 A; 239 C; 250 G; 307 T; 0 other;

Query Match 65.5%; Score 748.6; DB 20; Length 1147;
Best Local Similarity 79.8%; Pred. No. 3.2e-215;
Matches 883; Conservative 0; Mismatches 224; Indels 0; Gaps 0;

QY 37 TTCACCTGTTTCAGAGATGATGACATCAGAAATGGAATGTTCTTTCTCT 96
D 1 TTCGACTTATTCAGAGATGATGATGACCAAGGAATGGAATGTTCTTTCTTC 60
QY 97 CTGAGCCTCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTCTCT 156
D 61 CTGAGCATCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
QY 157 CAGATTGATAGTTGCTTCATGTTTAACTGCTGCTGCTGCTGCTGCTGCTGCT 216
D 121 CAGATTGCAAGGCACTGCTGCTTAACTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 217 CAGTCAGGCTCTCAGTCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 276
D 181 CAGCAGGACTCTGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 277 GATTATGATCTCAGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 336
D 241 GATTATGAACTCAGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 337 GACTACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 396
D 301 AACTACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 397 AATCATTTTGAAGACACTAGACGTAATTAATTAATTAATTAATTAATTAAT 456
D 361 AATGATGTAAGATACAGATCCAGATTTAAATTAATTAATTAATTAATTAAT 420
QY 457 AAAATCAAGAACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 516
D 421 AAGATCAAGAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 517 AATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 576
D 481 AAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 577 TGCATTTTCAATCTCCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 636
D 541 TSCCGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 637 AAGTCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 696
D 601 CGGTTCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 697 GTGGCATTAACATGTAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 756
D 661 GTGGCATTAACATGTAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 757 CTGACCTTTTCAATCTTAAATGTAAGTGTGCTGCTGCTGCTGCTGCTGCT 816
D 721 CTGAGTTTCCAGATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
QY 817 GAGGTATTTTCTCAGTCTAAGATGAGAAATTAATTAATTAATTAATTAAT 876
D 781 AAGCTGTTTCTCCCCCAGTTCAGATGAGAAATTAATTAATTAATTAATTAAT 840

```


Qy	1057	ACGCTGTTTAGAGCTGACACACCCATTCCTATTGTTATACAGAGAGATGACATCATCTTTA	1116
Db	1021	ACAGTGTTCAGAGCGCACCGGCCCTTCTGTTTGTTCATCAAGAGAGATGACATCATCTTTA	1080
Qy	1117	TTTCAGTGGCAAGTTTCTTGCCCTTGA	1143
Db	1081	TTTACTGGCAAGTCTCTTGTCCTTGA	1107
RESULT 12			
AAF82440			
ID	AAF82440	standard; cDNA; 1147 BP.	
XX	AAF82440;		
XX	10-JUL-2001	(first entry)	
XX	Mouse megin nucleotide sequence.		
DE			
XX	Mouse; megin; mesangial cell proliferative nephritis; nephrotropic;		
KW	transgenic mouse; glomerular disease; animal model; drug screening; ss.		
XX			
XX	Mus musculus.		
XX			
Key	Location/Qualifiers		
FH	1..1107		
FT	CDS		
FT	/*tag= a		
FT	/product= "megsin"		
XX			
PN	WO200124628-A1.		
XX			
PD	12-APR-2001.		
XX			
PF	06-OCT-2000; 2000WO-JP06988.		
XX			
PR	06-OCT-1999; 99JP-0285736.		
XX			
XX	(KURO/) KUROKAWA K.		
PA	(MIYA/) MIYATA T.		
XX			
PI	Miyata T;		
XX			
WPI	2001-300136/31.		
DR	P-PSDB; AAB83077.		
XX			
XX	Mouse model for mesangial cell proliferative nephritis for development		
FT	and screening of new treatments		
PT			
XX			
PS	Disclosure; Page 50-52; 62pp; Japanese.		
XX			
CC	The present sequence encodes mouse megin. The human megin coding		
CC	sequence may be introduced into a mouse to produce an animal model of		
CC	mesangial cell proliferative nephritis. The symptoms include		
CC	enlargement of the mesangial base region, sedimentation of an immune		
CC	complex and an increase in mesangial cells. The animal model is useful		
CC	for analysing the pathology of chronic glomerular diseases and for		
CC	screening compositions for prevention and treatment of the diseases.		
CC	highly uniform models can be made easily and in large numbers using		
CC	this method.		
XX			
SQ	Sequence 1147 BP; 351 A; 239 C; 250 G; 307 T; 0 other;		
Query Match	65.58; Score 748.6; DB 22; Length 1147;		
Best Local Similarity	79.8%; Pred.No.3.2e-215;		
Matches 883; Conservative	0; Mismatches 224; Indels 0; Gaps 0;		
Qy	37	TTCAAGCTCTTCAGAGATGATGACAAATCAAGAAATGGAATCTGTTCTTCTCTCT 96	
Db	1	TTGACTTTTTCAGAGATGATGATAGCCAGGAATGTAATCTTCTCTCTCC 60	
Qy	97	CTGAGCTCTTCGCTGGCCCTCGCCCTGGTCCGCTTGGGCGCTCAAGATGATCCCTCT 156	

ID ACA56226 standard; cDNA; 316 BP.

AC ACA56226;

DT 06-JUN-2003 (first entry)

DE Human signalling pathway polynucleotide probe SEQ ID NO 824.

Human; probe; ss; array element; Parkinson's disease;

signalling pathway population; cancer; adenocarcinoma; leukaemia;

immunopathy; AIDS; asthma; neuropathy; Alzheimer's disease; microarray.

OS Homo sapiens.

XX US6500938-B1.

XX 31-DEC-2002.

XX 30-JAN-1998; 98US-0016434.

XX 30-JAN-1998; 98US-0016434.

XX (INCY-) INCYTE GENOMICS INC.

XX Au-Young J, Seilhamer JJ;

XX WPI; 2003-352189/33.

XX Combination of polynucleotide probes, useful as array elements in a

microarray for monitoring the expression of a number of target

polynucleotides -

XX Claim 1; SEQ ID NO 824; 65pp; English.

XX The invention relates to a combination which, comprises a number of

polynucleotide probes comprising a sequence selected from one of the 1490

sequences mentioned in the specification. The combination is useful as an

array element in a microarray for monitoring the expression of a number

of target polynucleotides. The microarray is particularly useful in the

diagnosis and treatment of cancer and immunopathology and neuropathology.

The microarray is useful in diagnostics and treatment regimens, drug

discovery and development, toxicological and carcinogenicity studies,

forensics and pharmacogenomics. The microarray is also useful for

monitoring progression of diseases and for developing sophisticated

profiles for the effects of currently available therapeutic drugs. The

combination is also useful for purifying a subpopulation of mRNAs, cDNAs

and genomic fragments and in research and diagnostic applications. The

array can detect changes in expression in a large number of genes coding

for different signaling pathway populations which can be used to diagnose

various diseases including cancer e.g. adenocarcinoma and leukaemia,

immunopathies e.g. AIDS and asthma, neuropathies e.g. Alzheimer's disease

and Parkinson's disease. The present sequence represents a polynucleotide

probe of the invention.

Note: The sequence data for this patent did not form part of the printed

specification but was obtained in electronic format directly from USPTO

at seqdata.uspto.gov/sequence.html?DocID=06500938B1.

Query Match

Best Local Similarity 26.68; Score 304; DB 25; Length 316;

Matches 315; Conservative 0; Mismatches 1; Gaps 1;

822 ATTATTTCTCAGTTCAAGATAGAGAAGATTATGAATGAATGAATGAATTTGGAGCCCT 881

1 ATTATTTCTCAGTTCAAGATAGAGAAGATTATGAATGAATGAATTTGGAGCCCT 60

892 AGGGCTGAAGATATCTTTGATGAATCCAAAGCAGATCTCTGGGATTTCTCGGGGG 941

61 AGGGCTGAAGATATCTTTGATGAATCCAAAGCAGATCTCTGGGATTTCTCGGGGG 120

942 TCGTCTGATATATCAAGATGATCCAAATTTACATPAGAGTCTCAGGAGGGAC 1001

Db 121 TCGTCTGATATATCAAGATGATCCAAATTTACATAGAGTCACTGAGAGGGAC 180

Qy 1002 CGAGGCTACTGTGTCACAGGAAGTAAATTTCTAGAAAAGCAAACTCCCTCAGTCCAGCT 1061

Db 181 CGAGGCTACTGTGTCACAGGAAGTAAATTTCTAGAAAAGCAAACTCCCTCAGTCCAGCT 240

Qy 1062 GTTTAGAGCTGACACCCATCTCTATTTCTATCAGGAAGATGATCACTTATTTCAG 1121

Db 241 GTTTAGAGCTGACACCCATCTCTATTTCTATCAGGAAGT-TCATCATCTTATTTCAG 299

Qy 1122 TGGCAAAAGTTTCTGCC 1138

Db 300 TGGCAAAAGTTTCTGCC 316

RESULT 14

ACA56060

ID ACA56060 standard; cDNA; 254 BP.

XX AC ACA56060;

XX 06-JUN-2003 (first entry)

XX Human signalling pathway polynucleotide probe SEQ ID NO 658.

XX Human; probe; ss; array element; Parkinson's disease;

signalling pathway population; cancer; adenocarcinoma; leukaemia;

immunopathy; AIDS; asthma; neuropathy; Alzheimer's disease; microarray.

XX Homo sapiens.

XX US6500938-B1.

XX 31-DEC-2002.

XX 30-JAN-1998; 98US-0016434.

XX 30-JAN-1998; 98US-0016434.

XX (INCY-) INCYTE GENOMICS INC.

XX Au-Young J, Seilhamer JJ;

XX WPI; 2003-352189/33.

XX Combination of polynucleotide probes, useful as array elements in a

microarray for monitoring the expression of a number of target

polynucleotides -

XX Claim 1; SEQ ID NO 658; 65pp; English.

XX The invention relates to a combination which, comprises a number of

polynucleotide probes comprising a sequence selected from one of the 1490

sequences mentioned in the specification. The combination is useful as an

array element in a microarray for monitoring the expression of a number

of target polynucleotides. The microarray is particularly useful in the

diagnosis and treatment of cancer and immunopathology and neuropathology.

The microarray is useful in diagnostics and treatment regimens, drug

discovery and development, toxicological and carcinogenicity studies,

forensics and pharmacogenomics. The microarray is also useful for

monitoring progression of diseases and for developing sophisticated

profiles for the effects of currently available therapeutic drugs. The

combination is also useful for purifying a subpopulation of mRNAs, cDNAs

and genomic fragments and in research and diagnostic applications. The

array can detect changes in expression in a large number of genes coding

for different signaling pathway populations which can be used to diagnose

various diseases including cancer e.g. adenocarcinoma and leukaemia,

immunopathies e.g. AIDS and asthma, neuropathies e.g. Alzheimer's disease

and Parkinson's disease. The present sequence represents a polynucleotide

probe of the invention.

Note: The sequence data for this patent did not form part of the printed

specification but was obtained in electronic format directly from USPTO

at seqdata.uspto.gov/sequence.html?DocID=06500938B1.

```

XX SQ Sequence 254 BP; 83 A; 38 C; 64 G; 69 T; 0 other;
Query Match 22.2%; Score 254; DB 25; Length 254;
Best Local Similarity 100.0%; Pred. No. 2.9e-66;
Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 303 GCTTTTGTGAAAGGTGATGGCTTTTCATAGGACTACATTGAGTGTGCCGAAATTT 362
Dd |||||
1 GCTTTTGTGAAAGGTGATGGCTTTTCATAGGACTACATTGAGTGTGCCGAAATTT 60
QY 363 ATACGATGCCAAAGTGGAGCGAGTTGACTTTACGAATCATTTAGAGACACTAGACTAA 422
Dd |||||
61 ATACGATGCCAAAGTGGAGCGAGTTGACTTTACGAATCATTTAGAGACACTAGACTAA 120
QY 423 TATTAATAGTGGTTGAAATGAACACACATGGCAAAATCAGAACGTGATTTGGTGAAG 482
Dd |||||
121 TATTAATAGTGGTTGAAATGAACACACATGGCAAAATCAGAACGTGATTTGGTGAAG 180
QY 483 TGGCATAAGCTCATCTGCTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 542
Dd |||||
181 TGGCATAAGCTCATCTGCTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 543 GCAATCAGCTTCA 556
Dd |||||
241 GCAATCAGCTTCA 254
RESULT 15
ABA90345
ID ABA90345 standard; cDNA; 1185 BP.
XX
AC ABA90345;
XX
DT 12-FEB-2002 (first entry)
XX
DE Human polynucleotide #20.
XX
KW Human; neurotropic; neuroprotective; anticonvulsant; antidepressant;
KW neuroleptic; tranquiliser; antiarrhythmic; cardiac; antiasthmatic;
KW antiinflammatory; antilipemic; hepatotropic; virucide; antidiabetic;
KW nephrotropic; anorectic; cytostatic; vaccine; neurological disease;
KW cardiovascular disease; respiratory disease; liver disease;
KW renal disease; skeletal muscle disease; gastrointestinal disease;
KW placental disease; testicular cancer; male fertility; pancreatic disease;
KW ss.
XX
OS Homo sapiens.
XX
PN WO200181363-A1.
XX
PD 01-NOV-2001.
XX
PF 26-APR-2001; 2001WO-US13360.
XX
PR 27-APR-2000; 2000US-199963P.
PR 11-MAY-2000; 2000US-203336P.
PR 25-MAY-2000; 2000US-207087P.
PR 26-MAY-2000; 2000US-207546P.
XX
XX (SMIK) SMITHKLINE BEECHAM CORP.
XX (SMIK) SMITHKLINE BEECHAM PLC.
XX
XX Agarwal P, Murdoch PR, Rizvi SK, Smith RF, Xiang Z, Kabnick KS;
XX Lai Y, Xie Q;
XX
XX WPI; 2002-041392/05.
XX P-PSDB; ABB53280.
XX
XX Novel polypeptides and polynucleotides useful as a vaccine for
XX preventing and treating diseases associated with the polypeptide, e.g.
XX Alzheimer's disease, dyslipidemia, obesity, diabetes, infertility,
XX asthma, amnesia -

```

```

XX Claim 2; Page 56; 116pp; English.
XX
XX The invention relates to an isolated polypeptide comprising a 277, 480,
CC 583, 581, 628, 424, 638, 229, 310, 841, 241, 369, 382, 185, 586, 1026,
CC 844, 782, 262, 394, 471, 485, 286, 533, 495, 350, 619, 430, 482, 255,
CC 784, 252, 593, 477, 607, 781, 640, 686 or 154 amino acid sequence as
CC given in the specification. The polypeptides, modulators of the
CC treating diseases such as neurological and psychiatric diseases
CC including Alzheimer's, parkinsonism, Huntington's disease,
CC myotonic dystrophy, anorexia and depression; cardiovascular diseases,
CC including congestive heart failure, Hodgkin's disease and myocardial
CC infarction; respiratory diseases including asthma, chronic obstructive
CC pulmonary disease, cystic fibrosis and adult respiratory distress
CC syndrome; liver diseases including hypercholesterolaemia, cirrhosis,
CC viral and nonviral hepatitis, Type II diabetes mellitus, and impaired
CC glucose tolerance; renal disease including renal failure, acute tubular
CC necrosis and glomerulonephritis; skeletal muscle diseases including
CC Eulenburg's disease, hypoglycaemia and obesity; gastrointestinal
CC diseases including myotonia congenita and intestinal obstruction; lymph
CC diseases including lymphagiectasia; diseases of placenta including
CC choriocarcinoma; diseases of testes including testicular cancer,
CC male reproductive diseases including low testosterone and male
CC infertility; and disease of pancreas including diabetic ketoacidosis,
CC Type 1 and 2 diabetes and obesity. The present sequence encodes a
CC polypeptide of the invention.
XX
SQ Sequence 1185 BP; 355 A; 257 C; 268 G; 305 T; 0 other;
Query Match 21.3%; Score 243.6; DB 24; Length 1185;
Best Local Similarity 53.8%; Pred. No. 9.2e-63;
Matches 616; Conservative 0; Mismatches 494; Indels 36; Gaps 4;
QY 1 ATGCGCTCCCTTGTGCGAGCAATGCGAGATTTTGTCTTCAACCTCTTCAGAGAGATGGAT 60
Dd |||||
1 ATGCGCTCTCTGTGTACAGCAACACCAATTTTGTCTTGTATCTTTTCAAGAGATAGGC 60
QY 61 GACAACTCAGGAATGGAATGTGTCTTCTCTCTCTGAGCTCTTCTGCTGCTGCTGCTGCT 120
Dd |||||
61 AAAGATGATGCTGCTAATAAACAATATTTCTCTCTCTGAGCTCTCAGTGTGCTGCTGCT 120
QY 121 CTGCTCGCTTGGCGCTCAAGATGACTCCCTCTCTCAGATGATAAGTGTCTTCAATGTT 180
Dd |||||
121 ATGCTAGCTTGGGTGTAGAGTGACAGTGCACATCAGATTGAGGTACTACTACTTC 180
QY 181 A-----ACACTGCTCAGGATGGAATACTCTTCTAATAGTCAGTCAAGGCTCCAGT 232
Dd |||||
181 AACAAACGACAGAGCTCTGGATCAGCAGGCTGGTCTTAAACAATGAGAGCGAGCTG 240
QY 233 CTCAACTGAAA-----AGAGTTTCTGATATAATGATCCCAAGGATTTAGAT 285
Dd |||||
241 GTCACTGCTACTTTGGGCGAGCTCTCTCCAAATTAGACAGGATCAAGACTGATTACACA 300
QY 286 CTCACTGCTGTAATGGGCTTTTGTGAAAAGTGTATGCTTCTAATAAGGACTACTATT 345
Dd |||||
301 CTGATGATGCGACAGGCTTTATGAGAGCAGAAATCCCAATCTCTCAGGATTTACTTA 360
QY 346 GATGTGCGGAAAATTTATACGATGCCAAAGTGGAGCGAGTTGACTTTCAGATCATTTTA 405
Dd |||||
361 GATGTGTGATTTCAATTTTACACACAGCAGATTGAAAGTGTGATTTTCAAAAAACCCCT 420
QY 406 GAAGCACTGACGCTATATTAATTAAGTGGGTGAAAATGAAACACATGCAAAATCAAG 465
Dd |||||
421 GAAAAATCCACAGAGAGATTAACTTCTGGGTGAAATGTCAATCCCAAGGTAAATCAAG 480
QY 466 AACGTGATTTGTAAGGTGGCATAAGCTCTCTGTTAATGCTGCTGCTGCTGCTGCTG 525
Dd |||||
481 GAATCTTTCACAGAGAGAGCTATTATCTGAGACTGTGCTGCTGCTGCTGCTGCTGCT 540
QY 526 TACTTCAAGGCAAGTGGGAATTCAGCTCTTCAAGAGAGGCAAAACCAATAATGCCATTC 585
Dd |||||
541 TACTTCAAGGCAAAATGGGAACATATCTTTGACCATGAAACACGGTGGATGCACCTTTC 600

```

Search completed: December 14, 2003, 01:19:47
Job time : 384 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 13, 2003, 22:10:21 : Search time 89 Seconds
(without alignments)
5668.550 Million cell updates/sec

Title: US-09-508-997A-1
Perfect score: 1143
Sequence: 1 atgggcctccctgtgcagc.....gcaaacctttctgccttga 1143

Scoring table: IDENTITY_NUC
Gapop = 0.0 , Gapex: 1.0

Searched: 569978 segs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTOUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1143	100.0	1950	2	US-08-472-659-30
2	1143	100.0	1950	2	US-08-474-661-30
3	1143	100.0	1950	2	US-08-611-977-30
4	304	26.6	316	4	US-09-016-434-824
5	254	22.2	254	4	US-09-016-434-658
6	225.6	19.7	1512	1	US-07-768-286B-3
7	222	19.4	1482	1	US-07-768-286B-5
8	220.6	19.3	1308	4	US-09-123-912-109
9	220.6	19.3	1308	4	US-09-643-597-109
10	220.6	19.3	1308	4	US-09-480-884A-109
11	220.6	19.3	1308	4	US-09-542-615A-109
12	220.6	19.3	1308	4	US-09-606-421B-109
13	211.6	18.5	1711	1	US-08-568-147B-1
14	210	18.4	1245	4	US-09-266-910-1
15	210	18.4	1245	4	US-09-266-910-2
16	206.8	18.1	1328	1	US-07-679-052A-14
17	206.8	18.1	1340	1	US-07-679-052A-16
18	206.8	18.1	1610	1	US-07-768-286B-1
19	206.8	18.1	2409	1	US-07-911-531-18
20	206.8	18.1	2409	1	US-07-693-636A-18
21	205	17.9	1419	4	US-09-123-912-111
22	205	17.9	1419	4	US-09-643-597-111
23	205	17.9	1419	4	US-09-480-884A-111
24	205	17.9	1419	4	US-09-542-615A-111
25	205	17.9	1419	4	US-09-606-421B-111
26	188.4	16.5	1278	4	US-10-024-427-1
27	188.4	16.5	1519	4	US-10-024-427-3

28	186.4	16.3	1366	3	US-09-200-965-1	Sequence 1, Appli
29	185.6	16.2	1361	4	US-09-016-434-1284	Sequence 1284, Ap
30	182.4	16.0	1230	4	US-09-613-303-54	Sequence 54, Appl
31	182.4	16.0	1873	4	US-08-545-573A-38	Sequence 38, Appl
32	182.4	16.0	2847	4	US-09-613-303-20	Sequence 20, Appl
33	179.6	15.7	1425	1	US-08-464-148-1	Sequence 1, Appli
34	179.6	15.7	1425	1	US-08-365-500-1	Sequence 1, Appli
35	179.6	15.7	1425	1	US-08-846-784-1	Sequence 1, Appli
36	170	14.9	1393	1	US-09-016-434-1075	Sequence 1075, Ap
37	170	14.9	1393	1	US-08-464-148-3	Sequence 3, Appli
38	170	14.9	1393	1	US-08-365-500-3	Sequence 3, Appli
39	170	14.9	1393	1	US-08-846-784-3	Sequence 3, Appli
40	156.8	13.7	2584	1	US-08-121-714-1	Sequence 1, Appli
41	156.8	13.7	2584	1	US-08-322-742-13	Sequence 13, Appl
42	156.8	13.7	2584	1	US-08-477-108A-1	Sequence 1, Appli
43	156.8	13.7	2584	2	US-08-477-112-1	Sequence 1, Appli
44	156.8	13.7	2584	5	PCT-US93-08322-1	Sequence 1, Appli
45	155.6	13.6	1152	1	US-08-315-831A-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-08-472-659-30
; Sequence 30, Application US/08472659
; Patent No. 5831030
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Fuyuki
; APPLICANT: TSUBOUKKA, No. 5831030uo
; APPLICANT: NAKAZATO, Hiroshi
; APPLICANT: MIURA, Kenju
; APPLICANT: ISHIDA, No. 5831030uhito
; APPLICANT: KURIHARA, Tatsuya
; APPLICANT: YAMAUCHI, Kojo
; APPLICANT: YAMAGUCHI, No. 5831030omi
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,659
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5-067339
; FILING DATE: 04-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm X.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-248
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 1950 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; ORIGINAL SOURCE:
 ; ORGANISM: Homo sapiens
 ; STRAIN: A431
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 1
 ; OTHER INFORMATION: /note= "DNA coding for human
 ; OTHER INFORMATION: megakaryocyte differentiation factor."
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 74..1217
 ; US-08-472-659-30

Query Match 100.0%; Score 1143; DB 2; Length 1950;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGCGCTCCCTGCTGCAGCAAAATGCGAGATTTTTCCTCAACCTCTTCAGAGAGATGAT	60
Db	74	ATGCGCTCCCTGCTGCAGCAAAATGCGAGATTTTTCCTCAACCTCTTCAGAGAGATGAT	133
QY	61	GACATCAAGGAATGGAATGTGTCTTTTCCTCTCAGAGCTCTTCGCTGCCCTGGCC	120
Db	134	GACATCAAGGAATGGAATGTGTCTTTTCCTCTCAGAGCTCTTCGCTGCCCTGGCC	193
QY	121	CTGCTCCGCTGGGCGTCAAGATGACTCCCTCTCAGATTTGATAGTTCCTCATGTT	180
Db	194	CTGCTCCGCTGGGCGTCAAGATGACTCCCTCTCAGATTTGATAGTTCCTCATGTT	253
QY	181	AACATGCTCAGGATATGGAATCTTCTTAATAGTCAGTCAGGCTCCAGTCTCACTG	240
Db	254	AACATGCTCAGGATATGGAATCTTCTTAATAGTCAGTCAGGCTCCAGTCTCACTG	313
QY	241	AAAGAGTTTTTCTGATATTAATGATCCCAAGGATTAATGATCAGCATGTGAT	300
Db	314	AAAGAGTTTTTCTGATATTAATGATCCCAAGGATTAATGATCAGCATGTGAT	373
QY	301	GGCTTTTGTGAAAAGTGTATGCTTCTCAATAGGATGATGATGATGATGATGATG	360
Db	374	GGCTTTTGTGAAAAGTGTATGCTTCTCAATAGGATGATGATGATGATGATGATG	433
QY	361	TTATACGATCCCAAGTGGAGCGAGTGTACCTTTACGAATCATTTAGAGACACTAGC	420
Db	434	TTATACGATCCCAAGTGGAGCGAGTGTACCTTTACGAATCATTTAGAGACACTAGC	493
QY	421	AATATTAATAGTGGTTGAAATGAACATGGAACATGGAACATGGAACATGGAAC	480
Db	494	AATATTAATAGTGGTTGAAATGAACATGGAACATGGAACATGGAACATGGAAC	553
QY	481	GCTGGCATTAAGTCTATCTGCTGTAATGCTGTGCTGCTGCTGCTGCTGCTGCTG	540
Db	554	GCTGGCATTAAGTCTATCTGCTGTAATGCTGTGCTGCTGCTGCTGCTGCTGCTG	613
QY	541	TGGCAATACGCTTCCCAAGCGAAGCAATTAATGCAATTTCAATCTCCCAAGTGC	600
Db	614	TGGCAATACGCTTCCCAAGCGAAGCAATTAATGCAATTTCAATCTCCCAAGTGC	673
QY	601	TCGGGAAGCGATGCGCATGATGATGATGATGATGATGATGATGATGATGATGATG	660
Db	674	TCGGGAAGCGATGCGCATGATGATGATGATGATGATGATGATGATGATGATGATG	733
QY	661	GACCATCAATGAGATTTCTGAGCTCAGATACATGATGATGATGATGATGATGATG	720
Db	734	GACCATCAATGAGATTTCTGAGCTCAGATACATGATGATGATGATGATGATGATG	793
QY	721	CTGCTGAGATGACCTCTCGAAATGAAATGAAATGAAATGAAATGAAATGAAATG	780
Db	794	CTGCTGAGATGACCTCTCGAAATGAAATGAAATGAAATGAAATGAAATGAAATG	853

QY	791	TGGACCAATCCAGCGAATGACCTCTTAAGTATCTTGAAGTATTTTTCCTCAGTTCAAG	840
Db	854	TGGACCAATCCAGCGAATGACCTCTTAAGTATCTTGAAGTATTTTTCCTCAGTTCAAG	913
QY	841	ATAGAGAGAAATATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG	900
Db	914	ATAGAGAGAAATATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG	973
QY	901	GATGAATCCAAAGAGATCTCTCTGGGATTTGCTTCGGGGGGTCTGCTCTATATATCAAG	960
Db	974	GATGAATCCAAAGAGATCTCTCTGGGATTTGCTTCGGGGGGTCTGCTCTATATATCAAG	1033
QY	961	ATGATGCAAAATCTTACATAGAGTCTACTGAGAGGACCCGAGGCTACTGCTGCAACA	1020
Db	1034	ATGATGCAAAATCTTACATAGAGTCTACTGAGAGGACCCGAGGCTACTGCTGCAACA	1093
QY	1021	GGAAGTAATATGTAAGAAAGCACTCCCTCAGTCCACGCTTTTAGAGCTGACCAACCA	1080
Db	1094	GGAAGTAATATGTAAGAAAGCACTCCCTCAGTCCACGCTTTTAGAGCTGACCAACCA	1153
QY	1081	TTCTATTTTGTATCAGAAAGGATGACATCATCTTATTCAGTGGCAAAAGTTCTTGCCCT	1140
Db	1154	TTCTATTTTGTATCAGAAAGGATGACATCATCTTATTCAGTGGCAAAAGTTCTTGCCCT	1213
QY	1141	TGA 1143	
Db	1214	TGA 1216	

RESULT 2
 US-08-474-661-30
 ; Sequence 30, Application US/08474661
 ; Patent No. 5874253
 ; GENERAL INFORMATION:
 ; APPLICANT: TSUJIMOTO, Masafumi
 ; APPLICANT: IWASA, Fuyuki
 ; APPLICANT: TSURUOKA, No. 5874253uo
 ; APPLICANT: NAKAZATO, Hiroshi
 ; APPLICANT: MIURA, Kenju
 ; APPLICANT: ISHIDA, No. 5874253uhio
 ; APPLICANT: KURIHARA, Tatsuya
 ; APPLICANT: YAMAICHI, Kozo
 ; APPLICANT: YAMAGUCHI, No. 5874253omi
 ; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
 ; NUMBER OF SEQUENCES: 34
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Burns, Doane, Swecker & Mathis
 ; STREET: George Mason Bldg., Washington & Prince Sts.
 ; CITY: Alexandria
 ; STATE: Virginia
 ; COUNTRY: United States
 ; ZIP: 22313-1404
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/474,661
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/091,028
 ; FILING DATE: 14-JUL-1993
 ; APPLICATION NUMBER: JP 4-212305
 ; FILING DATE: 17-JUL-1992
 ; PRICE APPLICATION DATA:
 ; APPLICATION NUMBER: JP 5-067339
 ; FILING DATE: 04-MAR-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: REA, TERESA STANEK
 ; REGISTRATION NUMBER: 30,427

Matches	1143;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	ATGSCCTCCCTTCCTGCAGCAAAATGAGAGTTTGCTTCAACCTGTTCAGAGAGATGGAT	60						
Db	74	ATGSCCTCCCTTCCTGCAGCAAAATGAGAGTTTGCTTCAACCTGTTCAGAGAGATGGAT	133						
QY	61	GACAAATCAAGAAATGGAATGTCTTTTCTCTCTGAGCTCTTCGCTGCCTGCAC	120						
Db	134	GACAAATCAAGAAATGGAATGTCTTTTCTCTCTGAGCTCTTCGCTGCCTGCAC	193						
QY	121	CTGGTTCGGCTTGGGCGCTCAAGATGACTCCCTCTCTCAGATTGATAAGTTGCTTCATGTT	180						
Db	194	CTGGTTCGGCTTGGGCGCTCAAGATGACTCCCTCTCTCAGATTGATAAGTTGCTTCATGTT	253						
QY	181	AACACTGCCTCAGGATATGGAAATCTCTATATAGTCAGTCAGGCTCCAGTCTCACAATG	240						
Db	254	AACACTGCCTCAGGATATGGAAATCTCTATATAGTCAGTCAGGCTCCAGTCTCACAATG	313						
QY	241	AAAAGAGTTTTTCTGATATAAATGCAATCCCAAGGATTAATGATCTCAGCATTGTGAAT	300						
Db	314	AAAAGAGTTTTTCTGATATAAATGCAATCCCAAGGATTAATGATCTCAGCATTGTGAAT	373						
QY	301	GGGCTTTTTCCTGAAAAGTGTGGCTTTTCAATGAGCACTTCAATTGAGTGTCCGAAAAA	360						
Db	374	GGGCTTTTTCCTGAAAAGTGTGGCTTTTCAATGAGCACTTCAATTGAGTGTCCGAAAAA	433						
QY	361	TTATACGATCCCAAGTGGAGCGATTTGACTTTACGAATCATTTAGAAGACACTAGAGCT	420						
Db	434	TTATACGATCCCAAGTGGAGCGATTTGACTTTACGAATCATTTAGAAGACACTAGAGCT	493						
QY	421	AATATTAATTAAGTGGTTGAAATGAAACACATGGGAAATCAAGACGTTGATGGTGA	480						
Db	494	AATATTAATTAAGTGGTTGAAATGAAACACATGGGAAATCAAGACGTTGATGGTGA	553						
QY	481	GATGCATTAAGCTCAATCTGCTGTAATGCTGTGGTGAATGCTGTCTACTTCAAGAGCAAG	540						
Db	554	GATGCATTAAGCTCAATCTGCTGTAATGCTGTGGTGAATGCTGTCTACTTCAAGAGCAAG	613						
QY	541	TGGCAATCAGCTTCCAGAGCGGAAACCAATAATTGGCATTTCAATCTCCCAAGTGC	600						
Db	614	TGGCAATCAGCTTCCAGAGCGGAAACCAATAATTGGCATTTCAATCTCCCAAGTGC	673						
QY	601	TCCTGGAGAGCTGCGCCATGTCATCAGGAAACGGAGTTCAATTTGCTGTATTATGAG	660						
Db	674	TCCTGGAGAGCTGCGCCATGTCATCAGGAAACGGAGTTCAATTTGCTGTATTATGAG	733						
QY	661	GACCATCAATGAAGATTTCTTTGAGCTCAGATACAAATGGTGGCATAAACATGTACGTTCTG	720						


```

/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Zeller, Karen J.
/ REGISTRATION NUMBER: 37,071
/ REFERENCE/DOCKET NUMBER: PA-0002 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (650) 855-0555
/ TELEFAX: (650) 845-4166
/ INFORMATION FOR SEQ ID NO: 658:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 254 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: XERANOT02
/ CLONE: 2622817
/ US-09-016-434-658

Query Match      22.2%; Score 254; DB 4; Length 254;
Best Local Similarity 100.0%; Pred. No. 7.7e-73;
Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY   303  GCCTTTTGGTCGAAAAGTGATAGCGCTTTCAATAAGACTACATTGAAGTGTGCCGAAAAAATT 362
Db    1  GCCTTTTGGTCGAAAAGTGATAGCGCTTTCAATAAGACTACATTGAAGTGTGCCGAAAAAATT 60

QY   363  ATACGATGCCAAAGTGGAGCGAGTTGCATTTACGAATCATTTAGAAGACACTAGACGTAA 422
Db    61  ATAAGATGCCAAAGTGGAGCGAGTTGCATTTACGAATCATTTAGAAGACACTAGACGTAA 120

QY   423  TATTAAATAAGTGGGTGTGAAAATGAAPACACATGCGCAAAATCAAGACGTGATGGTGAAGG 482
Db   121  TATTAAATAAGTGGGTGTGAAAATGAACAACATGCGCAAAATCAAGACGTGATGGTGAAGG 180

QY   483  TGGCATATAGCTCATCTGCTGTATATGCTGTGTGATGCTGTGTAAGTCTGTGTAAGGCAAGTG 542
Db   181  TGGCATATAGCTCATCTGCTGTATATGCTGTGTGTAAGTCTGTGTAAGGCAAGTG 240

QY   543  GCAATCAGCCCTTCA 556
Db   241  GCATCAGCCCTTCA 254

RESULT 6
US-07-768-286B-3
/ Sequence 3, Application US/07768286B
/ Patent No. 5444153
/ GENERAL INFORMATION:
/ APPLICANT: GOSS, Neil H.
/ APPLICANT: RICHARDSON, Michael A.
/ TITLE OF INVENTION: VARIANTS OF PAL-2
/ NUMBER OF SEQUENCES: 21
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Foley & Lardner
/ STREET: 1800 Diagonal Road, Suite 500
/ CITY: Alexandria
/ STATE: Virginia
/ COUNTRY: USA
/ ZIP: 22313-0299
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07768,286B
/ FILING DATE: 19911011
/ CLASSIFICATION: 514
/ PRIOR APPLICATION DATA:

```


RESULT 7
US-07-768-286B-5
; Sequence 5, Application US/07768286B
; Patent No. 544153
; GENERAL INFORMATION:
; APPLICANT: GOSS, Neil H.
; APPLICANT: RICHARDSON, Michael A.
; TITLE OF INVENTION: VARIANTS OF PAI-2
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 223-3-0289
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/768,286B
 FILING DATE: 19911011
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/AU90/00603
 FILING DATE: 20-DEC-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: BENT, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 16786/157 CHNC
 TELEPHONE: (703)836-9300
 TELEFAX: (703)683-4109
 TELEX: 899149
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1482 base pairs
 TYPE: NUCLEIC ACID
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 CELL TYPE: Monocyte
 CELL LINE: U937
 IMMEDIATE SOURCE:
 CLONE: BTA 1922
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 22..1170
 IDENTIFICATION METHOD: experimental
 OTHER INFORMATION: /function= "Product binds to
 OTHER INFORMATION: urokinase, tissue plasminogen activator"
 OTHER INFORMATION: /product= "PAI-2 variant, protease sensitive site
 OTHER INFORMATION: removed"
 OTHER INFORMATION: /evidence= EXPERIMENTAL
 OTHER INFORMATION: /note= "Codes for human plasminogen activator
 OTHER INFORMATION: inhibitor type 2 protein in which amino acids 66
 OTHER INFORMATION: to 98 inclusive have been deleted."
 US-07-768-286B-5

Query Match 19.4%; Score 222; DB 1; Length 1482;
 Best Local Similarity 52.2%; Pred. No. 6.5e-62;
 Matches 591; Conservative 0; Mismatches 505; Indels 36; Gaps 3;

QY 1 ATGGCTCCCTTGCTGCGAGCAATGCGAGTTTGGCTTCACTCTTCAGAGATGGAT 60
 DB 22 ATGGAGGATCTTTGTGGCAACACACTCTTGGCCCTCAATTTATTCAAGCATCTGGCA 81
 QY 61 GACAATCAAGGAATGGAATGTGTCTTCTCTCTGAGCTCTTCGCTGCCCTGGCC 120
 DB 82 AAAGCAAGCCCAAGCACTCTCTCTCTCCCATGGAGCATCTGCTCAGCATGGCC 141
 QY 121 CTGTCGCTGGGGCTCAAGATGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
 DB 142 ATGGTCTACATGGGCTCCAGGGGCGAGCAAGCAAGCATGGCCAGGCTGCTCAGTT 201
 QY 181 AACATGCTCTGAGTATGAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
 DB 202 ATGGAAGTGGAGCGCTGAGCA-----TAAATCATCTCATCTCTCTCTCTCTCT 250
 QY 241 AAAAGAGTTTTTCTGATATAAATGATCCCAAGGATATGATCTCAGCATGFGAAT 300
 DB 251 -----TCAGCTCTGCAATCAATGATCCACAGGGAATTTTCTGGAAGTGTCAAT 303
 QY 301 GGCTTTTGTGAAAAGTGTATGGCTTTTCATAGGACTACATTTAGTGTGCGGAAAAA 360
 DB 304 AAGCTTTTGTGAGAGTCTGGAGCTTCGGGAAGAAATATATCTGACTCTCTCAGAA 363
 QY 361 TTATAGATGCAAGTGGAGCGATTCATTTACGATCATTTAGAGACATGACGT 420

DB 364 TATTACTCTCTCAGAACCCCGCAGTAGACTTCTTAGAATGTCGACAGAGACTAGAAA 423
 QY 421 ATATTATTAAGTGGTGTGAAATGAAACACATGCGCAAAATCAAGAACCTGATTGGTGA 480
 DB 424 AAGATTAAATCTCTGGTCAAGACTCAACACAAAGGCAAAATCCCAAACTTGTACCTGAA 483
 QY 481 GGTGGCATAGCTCATCTGCTGTAATGGTCTGCTGGTGAATGCTGTGTACTTCAAGGCAAG 540
 DB 484 GTTCTCTAGATGGGATACCGAGATGGTCTGCTGGTGAATGCTGTGTACTTCAAGGAAAG 543
 QY 541 TGGCAATCAGCTTCCACAGAGCGAAACCAATAAATTGCCATTTTCAAAATCTCCCAAGTGC 600
 DB 544 TGGAAATCTCATTTGAGAGAACTAAATGGCTTTATCCTTTCGTTAAACTCGGCT 603
 QY 601 TCTGGAGGCGATCGCCATGATGCATCAAGAGCGAGTCAATTTGCTGTATTGAG 660
 DB 604 CAGCGCACCTGTACAGATGATGACTTGGCTGAAAGCTAAACATTGGATACATAGAA 663
 QY 661 GACCCATCAATGAAGATTTCTGAGCTCAGATACATATGCTGCATATAAACAATGTAGCTTCTG 720
 DB 664 GACCTAAAGGCTCAGATTTCTAGACTCCCATATGCTGGAGATGTAGCATGTTCTTGTG 723
 QY 721 CTGCTCA-----GAATGACCTCTCTGAAATTCGAAATCAAAACAACATCAACC 762
 DB 724 CTTCAGATGAAATGCGCATGTGTCCACTGGCTTGGAGCTGCTGAAAGTGAATAAACC 783
 QY 763 TTTCAGATCTAATGGAATGACCAATCCAGGCGGATGACCTCTAAGTATGTTGAGTGA 822
 DB 784 TATGACAACTCAACAGTGACCCAGCAAGACAAATGCTGAAGATGAATGAGTGAATGAGGTA 843
 QY 823 TTTTCTCTCAGTTCAAGATAGAGAAATATGAAATGAAACAAATATTTGAGAGCCCTA 882
 DB 844 TACATACCCAGTTCAMATTAGAGAGCATATGAACTCAGATCCATCTGAGAGCATG 903
 QY 883 GGGCTGAAAGATATCTTTGATGAATCCAAAGCAGATCTCTCTGGGATGCTTGGGGGGGT 942
 DB 904 GGCATGGAGAGCGCTTCAAAAGGAGCGGCCAAATTTCTCAGGATGTGCGAGAGGAA 963
 QY 943 CGTCTGATATATCAAGGATGATGCACAAATCTTACATAGAGTCACTCAGAGGAGGCAAC 1002
 DB 964 GACCTGTTCTTTCTGAAAGTGTCCACAGCCATGCTGTGATGTGAATGAGAGGAGGCACT 1023
 QY 1003 GAGCTACTGTCGCCACAGGAAGTAAATATTTAGAAAAGCACTCCCTCAGTCCACGCTG 1062
 DB 1024 GAAGACGCGCTGGCAAGAGGTGTATGACAGGGAGAACTGGACATGGAGGCGCCACAG 1083
 QY 1063 TTAGAGCTGACCCACCATCTCTATTTGTTATCAGGAAGGATGACATCATCT 1114
 DB 1084 TTTGTGGCAGATCATCTCTTTCTTTTCTTTTATTATGATTAAGATAACCAACT 1135

RESULT 8

US-09-123-912-109
 ; Sequence 109, Application US/09123912A
 ; Patent No. 6312695
 ; GENERAL INFORMATION:
 ; APPLICANT: Reed, Steven G.
 ; APPLICANT: Wang, Tongrong
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
 ; FILE REFERENCE: 210121.455C1
 ; CURRENT APPLICATION NUMBER: US/09/123,912A
 ; CURRENT FILING DATE: 1998-07-27
 ; PRIOR APPLICATION NUMBER: 09/040,802
 ; PRIOR FILING DATE: 1998-03-18
 ; NUMBER OF SEQ ID NOS: 114
 ; SOFTWARE: PatentIn ver. 2.0
 ; SEQ ID NO 109
 ; LENGTH: 1308
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-123-912-109

Query Match

19.3%; Score 220.6; DB 4; Length 1308;

```
Best Local Similarity 52.7%; Pred. No. 1.7e-61;
Matches 621; Conservative 0; Mismatches 519; Indels 39; Gaps 5;

QY 1 ATGGCTCCCTGCTGCGAAGAAATGAGAGATTTGCTTCAACCTGTTGAGAGAGATGAT 60
DB 25 ATGGATTCATTGGGCGCTGAGCACTGCGATTTGGTTGATCTTTTCAAGAGCTGAAG 84
QY 61 GACATCAAGGAATGAAATGTTCTTTTCTCTGAGCCTCTGCGTCCCTGGCC 120
DB 85 AAAACAATG---ATGGCAACATCTCTTTTCCCTGTGGGCATCTTGACTGCAATGGC 141
QY 121 CTGCTCCCTGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
DB 142 ATGGCTCCCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 201
QY 181 AACATGCTCTCA-----GGATATGGAATCTCTCTTAATAGTCAAG 219
DB 202 GAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 261
QY 220 TCAGGCTCCAGTCTCACTGAAAGAGATTTTCTGATTAATGATGATGATGATGAT 279
DB 262 GAAGCAGTACATCAACAAATTCACAAATTTTTCAGTGAATTAAGCAATCACTAAT 321
QY 280 TATGATCTCAGCATTTGTAATGGGCTTTTCTGAAAGATGATGATGATGATGATGAT 339
DB 322 TATGATCTGACATTAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 381
QY 340 TACATTTGATGCTGCGGAAATTTATGATGATGATGATGATGATGATGATGATGAT 399
DB 382 TACTTAGATTTATTTGAAATATTTATCATGATCTCTGAGAGCTTTGATTTTGAAT 441
QY 400 CATTTAGAGAGACACTAGAGCTATTAATTAATGATGATGATGATGATGATGATGAT 459

US-09-643-597-109
; Sequence 109, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643.597
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 109
; LENGTH: 1308
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-643-597-109

Query Match 19.3%; Score 220.6; DB 4; Length 1308;
Best Local Similarity 52.7%; Pred. No. 1.7e-61;
Matches 621; Conservative 0; Mismatches 519; Indels 39; Gaps 5;

QY 1 ATGGCTCCCTGCTGAGCAAAATGAGAGATTTTGTCTCAACCTGTTTCAAGAGATGAT 60
DB 25 ATGGATTCATTGGGCGCTGAGCACTGCGATTTGGTTGATCTTTTCAAGAGCTGAAG 84
QY 61 GACATCAAGGAATGAAATGTTCTTTTCTCTGAGCCTCTGCGTCCCTGGCC 120
DB 85 AAAACAATG---ATGGCAACATCTCTTTTCCCTGTGGGCATCTTGACTGCAATGGC 141
QY 121 CTGCTCCCTGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
DB 142 ATGGCTCCCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 201
QY 181 AACATGCTCTCA-----GGATATGGAATCTCTCTTAATAGTCAAG 219
DB 202 GAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 261
QY 220 TCAGGCTCCAGTCTCACTGAAAGAGATTTTCTGATTAATGATGATGATGATGAT 279
DB 262 GAAGCAGTACATCAACAAATTCACAAATTTTTCAGTGAATTAAGCAATCACTAAT 321
QY 280 TATGATCTCAGCATTTGTAATGGGCTTTTCTGAAAGATGATGATGATGATGATGAT 339
DB 322 TATGATCTGACATTAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 381
QY 340 TACATTTGATGCTGCGGAAATTTATGATGATGATGATGATGATGATGATGATGAT 399
DB 382 TACTTAGATTTATTTGAAATATTTATCATGATCTCTGAGAGCTTTGATTTTGAAT 441
QY 400 CATTTAGAGAGACACTAGAGCTATTAATTAATGATGATGATGATGATGATGATGAT 459
```

Db 442 GCAGCGAGTAAAGTGGAGAGAGATTAATTCCTGGGTGAGAGCAAAACAAATGAAGA 501
 QY 460 ATCAAGACGATGATGTGAGGTGGGATAGCTCATCTGCTGTAATGCTGCTGGTGAAT 519
 Db 502 ATCAAGGACCTGTTCCAGATGGCTCTATTAGTAGCTTACCAAGCTGGTGGTGAAC 561
 QY 520 GCTGTGATCTCAAGGCAAGTGGCAATCGACCTTACCAGAGCGAAACCAATATTC 579
 Db 562 ATGGTTTATTTAAAGGCAATGGGACAGGAGTTTAAAGAAAGAAATATAAGGAAG 621
 QY 580 CATTTCAATCTCCAGAGTCTCTGGGAGGCAAGTCCCATGATGATGATGATGATGATG 639
 Db 622 AAATTTGGATGAATAAGACACAAAGTAAATCTGTACAGATGATGATGATGATGATG 681
 QY 640 TTCAATTTGCTGTATTTAGAGACCAATCAATGAAGATTTCTGAGCTGATGAT---ACAA 696
 Db 682 TTTAGCTTACATTTCTGGAGGACTTGGAGGCAAAATTTCTAGGAGTTCATATATAAAC 741
 QY 697 GTGGCATAAACATGTACGTTCTGCTGCTGA-----GAATGACCTCTCTGAATGAA 750
 Db 742 AACGACCTAAGCATGTTTGTGCTTCTGCCCAACGATGATGATGATGATGATGATGAT 801
 QY 751 AACAACTGACCTTTCAGATCTTAATGGAATGGAATGGAATGGAATGGAATGGAATG 810
 Db 802 GATATAAATAAGTCTCGAGAAATTTGGTAGAGTGGACTAGTCCAGGSCATATGGAAGA 861
 QY 811 TATGTTGAGGTATTTTTCTCAGTTTCAAGATGAGAGATGAGAGATGAGAGATGAGAG 870
 Db 862 AAGGTGAATCTGCACTTGGCCCGTTTGGAGTGGAGACAGTTACGATCTAGAGCGGTC 921
 QY 871 TTGAGAGCCCTAGGCTGAAAGATATCTTTGATGAATCCAAAGCAGATCTCTCTGGAAT 930
 Db 922 CTGGTCCATCGGATGGGATGGGATGCTTCAATGAGCACAAGCCGACTACTCGGAATG 981
 QY 931 GCTTGGGGGTGCTGTGTATATCAAGAGATGATGATGATGATGATGATGATGATGATG 990
 Db 982 TGTGAGGCTCGGGTGTACGCCAGAGTTCTGACAGTTCTTGTGGCAGTAACT 1041
 QY 991 GAGGAGGCGCAGGCTACTGCTGCACAGAGATGATGATGATGATGATGATGATGATGAT 1050
 Db 1042 GAGGAGGCGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1101
 QY 1051 CAGTCCAGCTGTTTAGAGCTGACCAACCTATCTTTGTTATCAGAGAGTGTG----- 1105
 Db 1102 GGTATGAAATGTTCACTCCATCATCCCTTCTGTTCTTCTATCAGCACAATGATGCC 1161
 QY 1106 -ACATCATCTTATCAGTGGCAAGTTTCTTGGCCCTTGA 1143
 Db 1162 AACAGCATCTCTCTTCGGGAGATTTTCTTCTCCNTAA 1200

RESULT 10
 US-09-480-884A-109
 ; Sequence 109, Application US/09480884A
 ; Patent No. 6492597
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Fan, Liqun
 ; APPLICANT: Hosken, Nancy A.
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Fang, Gary R.
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.45506
 ; CURRENT APPLICATION NUMBER: US/09/480.884A
 ; CURRENT FILING DATE: 2001-08-27
 ; NUMBER OF SEQ ID NOS: 330
 ; SOFTWARE: FastSeq for Windows Version: 3.0
 ; SEQ ID NO 109
 ; LENGTH: 1308
 ; TYPE: DNA
 ; ORGANISM: Homo sapien

US-09-480-884A-109
 Query Match 19.3%; Score 220.6; DB 4; Length 1308;
 Best Local Similarity 52.7%; Pred. No. 1.7e-61;
 Matches 621; Conservative 0; Mismatches 519; Indels 39; Gaps 5;
 QY 1 ATGCGCTCCCTGCTGAGCAAAATGAGAGTGGTTCCTCAACTCTTCAGAGAGATGAT 50
 Db 25 ATGGATTCATTTGGCGCGTCAGCACTGAGCTTGGGTTGATCTTTTCAAGAGCTGAAG 84
 QY 61 GACAAATCAAGAGAAATGGAAATGTGTTTCTCTCTGAGCTCTCTGCGCCCTGGCG 120
 Db 85 ABAACAATG---ATGSCACAATCTTCTTCCCTCTGGGATCTTCACTGCAATGGC 141
 QY 121 CTGTCGCTGGGCGCTCAAGATGATGCTCCCTCTCTCAGATTGATAATGCTCTCATGTT 180
 Db 142 ATGTCCTCTCTGGGAGCCCGAGGAGCCCGCTTCCAGTGGAGGAGGTCTTTCACTCT 201
 QY 181 AACACTGCGCTCA-----GATATGGAACCTCTTCTAATAGTCAG 219
 Db 202 GAATAAGAGAGAGAGAGCTCAAGATTAAGAGTGAAGAAAGAGTGAATGAGACACA 261
 QY 220 TCAGGCTCCAGTCTCAACTGAAAGAGTGTTCCTCATATAAATGCATCCCAAGAGAT 279
 Db 262 GAAGCAGTACATCAACAATTCGAAAGTGTTCGACTGAATTAAGCAAACTCACTAATGAT 321
 QY 280 TATGATCTCAGATTTGTAATGGGCTTTTGTGTAAGAGTGTATGGCTTTCATAGGAC 339
 Db 322 TATGAATCAACATAACCAACAGCTGTTCGAGAGAAACATACCTCTCTTCAAAA 381
 QY 340 TACATGAGTGGCGGAAATTAATACGATGCCAAAGTGGAGGAGTGTGATTCAGAT 399
 Db 382 TACTAGATTAATGTTGAATAATTAATCATGATCTCTGSAACCTGTTGATTTGTAAT 441
 QY 400 CATTTAGAAGACACATAGACGTAATTAATTAATAGTGGCTTGAATAAGCAACATGSCAA 459
 Db 442 GCAGCGATGAAGTTCGAAAGAGATTAATTCCTGGGTTGAAAGCAAAACAAATGAAGA 501
 QY 460 ATCAAGACGATGATGTGAGGTGGCATAGCTCATCTGCTGTAATGCTGCTGGTGAAT 519
 Db 502 ATCAAGGACTTGTTCGAGATGGCTCTATTAGTAGCTCTACCAAGCTGGTGGTGAAC 561
 QY 520 GCTGTGATCTCAAGCAAGTGGCAATCAGCTTCCAGAGAGGAAACCAATATTCG 579
 Db 562 ATGTTTATTTTAAAGGCAATGGGACAGGAGTTTGAAGAAAGAAATATAAGAGAGAG 621
 QY 580 CATTTCAATCTCCCAAGTCTCTGGAGGCTCGCCATGATGATGATGATGATGATGATG 639
 Db 622 AAATTTGGATGAATAAGACACAAGTAAATCTGTACAGATGATGATGATGATGATGATG 681
 QY 640 TTCAATTTGCTGTATTTGAGAGCCCTCAATGAAGATTTCTGAGCTCAGAT---ACAA 696
 Db 682 TTTAGCTTACATTTCTGGAGGACTTGGAGGCAAAATTTCTAGGATTTCCATATAAAG 741
 QY 697 GTGGCATAAACATGTACGTTCTGCTGCTGA-----GAATGACCTCTCTGAATGAA 750
 Db 742 AACGACCTAAGCATGTTTGTGCTTCTGCCCAACGATGATGATGATGATGATGATGATG 801
 QY 751 AACAACTGACCTTTCAGATCTTAATGGAATGGAATGGAATGGAATGGAATGGAATG 810
 Db 802 GATATAAATAAGTCTCGAGAAATTTGGTAGAGTGGACTAGTCCAGGSCATATGGAAGA 861
 QY 811 TATGTTGAGGTATTTTTCTCAGTTTCAAGATGAGAGATGAGAGATGAGAGATGAGAG 870
 Db 862 AAGGTGAATCTGCACTTGGCCCGTTTGGAGTGGAGACAGTTACGATCTAGAGCGGTC 921
 QY 871 TTGAGAGCCCTAGGCTGAAAGATATCTTTGATGAATCCAAAGCAGATCTCTCTGGAAT 930
 Db 922 CTGGTCCATCGGATGGGATGGGATGCTTCAATGAGCACAAGCCGACTACTCGGAATG 981
 QY 931 GCTTGGGGGTGCTGTGTATATCAAGAGATGATGATGATGATGATGATGATGATGATG 990
 Db 982 TGTGAGGCTCGGGTGTACGCCAGAGTTCTGACAGTTCTTGTGGCAGTAACT 1041

TYPE: DNA
ORGANISM: Homo sapien
US-09-606-421B-109

Query Match 19.3%; Score 220.6; DB 4; Length 1308;
Best Local Similarity 52.7%; Pred. No. 1.7e-61;
Matches 521; Conservative 0; Mismatches 519; Indels 39; Gaps 5;

QY 1 ATGGCTCCCTGCTGTCAGCAATGACAGGTTTGTCTCAACCTGTCAGAGAGTGAAT 60
DB 25 ATGGATTCTCTGGCGCGTCAGCACTCGACTTGGGTTTGCATCTTTCAAGAGCTGAAG 84
QY 61 GACATCAAGAGAAATGAAATGTGTTCTTCTCTCTGAGCTCTTCGCTGCCCTGGCC 120
DB 85 AAAACAATG--ATGCAACATCTCTTTTCCCTGTGGCACTCTGATGCAATGGC 141
QY 121 CTGCTCCGCTTGGCGCTCAAGATGCTCCTCTCTCTCAAGTATGATAAGTTGCTTCATGTT 180
DB 142 ATGGTCTCTCTGGGACCCGAGGAGCCACCGCTTCCCACTTGGAGAGGTGTTTCACTCT 201
QY 181 AACACTGGCTCA-----GGATNTGGAATCTCTTANATGATGAG 219
DB 202 GAAAAGAGAGAGAGCTCAAGATTAAGCTGAGAAAAGAGGTCATGAGAACACA 261
QY 220 TCAGGCTCCAGTCTCAACTGAAAGAGTTTTTCTGATATAATGATGCCACAGGAT 279
DB 262 GAAGCAGTACATCAACAATTCGAAAGTTTTTGTGCTGAAATTAAGCAACTCACTATGAT 321
QY 280 TATGATCTCAGCATGTGAATGGCTTTTCTGTAAGAGTGTATGCTTTCTATAGGAC 339
DB 322 TATGAACTGAACATAACCAACAGGCTGTTTGGAGAAACACATACCTCTCTTCAAAA 381
QY 340 TACATGATGTGCGGAAAAATTTATACGATGCCAAAGTGGAGGTGATGCTTACGAT 399
DB 382 TACTTGATGATGTTGAAAAATTTATCATGATCTCTGGAACCTGTGATTTGTAAT 441
QY 400 CATTTAGAGACACTAGAGCTAATATTAATAGTGGGTTGAAATGAAACACATGCCAA 459
DB 442 GCAGCGATGAAGTCGAAAGAGATTTANTCTGGSTTGAAGCAAAACATGAAAA 501
QY 460 ATCAAGACGTGATGTGAGGTGGCATAGCTCATCTGCTGTAATGCTGTGCTGAT 519
DB 502 ATCAAGGACCTTGTCCAGATGCTCTATTAGTGTCTTACCAAGCTGTGCTGTGATC 561
QY 520 GCTGTGACTCTCAAGCAAGTGGCAATCAAGCTTCCCAAGAGCAAAACCAATATGC 579
DB 562 ATGGTTATTTTAAGGCAATGGCAGAGGTTTAAAGAGAAATATCTAGAGAGAG 621
QY 580 CATTTCAATCTCCAGTCTCTCGGAGGCAAGTCCCATGATGCTATCAGGAACGGAAG 639
DB 622 AAATTTTGGATGAATAGAGCAACAAGTAAATCTGTACATGATGACACAGAGCCATTCC 681
QY 640 TTCAATTTGCTGTATTGAGGACCCATCATGAAATCTTGTAGCTCAAT---ACAAT 696
DB 682 TTATGCTTCACTTTCTGGAGGACTTGAGGCCAAATTTCTAGGGATTCATATAAAAC 741
QY 697 GTGGCATAAATCATGTACGTTCTGCTGCTCA-----GAATGACCTCTCTGAATTGAA 750
DB 742 AACGACCTTAACATGTTTGTCTTCCCAAGCATGATGCTGGGAGAGATATA 801
QY 751 AACAACTGACCTTTCAGATCTAATGAATGAGCAATTCAGAGCAATGACCTTAAG 810
DB 802 GATAAAATAGTCTGAGAAATTTGGTAGTGGACTAGTCCAGGGCATATGGAAGAAAGA 861
QY 811 TATGTTGAGGATTTTCTCTGATGATGAGTGAAGAAATTAATGAATGAACATAT 870
DB 862 AAGTGATCTGCTGCTTGGCCCGTTTGGGTGGAGGACAGTACGATCTAGAGCGGTC 921
QY 871 TTGAGAGCCCTAGGCTGAAAGATATCTTTGATGAATCCAAAGCAGATCTCTGGGAT 930
DB 922 CTGGTGCATCGGGATGGGCGATGCTTCTAGTGACCAAAAGCCGCTACTCTGGGATG 981
QY 931 GCTTCGGGGGTGCTGTGTATATCAAGGATGATGCAAAATCTTACATAGAGTCACT 990

DB 982 TCCTCAGGCTCCGGTGTGTAGCCGCCAGAAATTCCTGACAGTTCCTTGTGGCAGTAACT 1041
QY 991 GAGGAGGACCGAGGCTTACTGTGCGCCAGGAAGTAAATTTGTAGAAAGCAACTCCCT 1050
DB 1042 GAGGAGGACCGAGGCTTACTGTGCGCCAGGCTGCACTGCGATAGGCTTTACTGTACATCCGCCCA 1101
QY 1051 CAGTCCACAGCTTTTAGAGCTGACCAACCTTCCTTCTTATCAGGAAGGATG----- 1105
DB 1102 GGTCAATGAAATGTTCACTGCAATCATCCTTCTCTGTTCTTATCAGGACCAATGAATCC 1161
QY 1106 -ACATCATCTTATCAGTGGCAAGTTTCTTGGCCCTTGA 1143
DB 1162 AACAGATCTCTTCTTGGCAGATTTTCTTCTCTTAA 1200

RESULT 13
US-08-568-147B-1
; Sequence 1, Application US/08568147B
; Patent No. 5783422
; GENERAL INFORMATION:
; APPLICANT: Suminami, Yoshinori
; APPLICANT: Kato, Hiroshi
; APPLICANT: Sekiguchi, Kiyoshi
; APPLICANT: Takeda, Katsumichi
; TITLE OF INVENTION: DNA FRAGMENT CODING FOR SQUAMOUS CELL
; TITLE OF INVENTION: CARCINOMA-ASSOCIATED ANTIGEN
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/568,147B
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 800,952
; FILING DATE: 02-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8425
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1711 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 62..1234
; US-08-568-147B-1

Query Match 18.5%; Score 211.6; DB 1; Length 1711;
Best Local Similarity 54.5%; Pred. No. 1.8e-58;
Matches 571; Conservative 0; Mismatches 444; Indels 33; Gaps 6;
QY 1 ATGGCTCCCTGCTGTCAGCAATGACAGGTTTGTCTCAACCTGTCAGAGAGTGAAT 60

Db	62	ATGAATTCACTCAGTGAAGACAACAACAAAGTTCATGTTGCAGCTGTTCACAGATTGAGA	121
QY	61	GACAAATCAAGGAAATGGAAATGTGTCTTTTCTCTCTGAGCTCTTCGCTGCCCTGCC	120
Db	122	AAATCAAAAGAGAA--CAACATCTCTTATTTCCCTATCAGCATCACATCAGCATTAGG	178
QY	121	CTGGTCGCTTGGGCGCTCAGATGACTCCCTCTCTCAGATTGATTAAGTTGCTTCATGTT	180
Db	179	ATGGTCTCTTAGAGCCCAAGACAACACATGCACAACGATTAAGAAGTCTTTCACATT	238
QY	181	-----AACACTGCTCAGGATGTAAGAACTCTCTTAATAGTCAGTCAGG--	224
Db	239	GATCAAGTCACAGAGAAACCAACAGGAAAGCTGCACATATCAATGTTGATAGTCAGGA	258
QY	225	--GCTCCAGTCTCAACTGAAAGAGTTTTTTCGATATAAATGCAATCCACAGAGTTAT	282
Db	299	AATGTTCAATCACAGATTTCAAAAGCTTCTGACTGAATTCACAACAATCCATGATCATAT	358
QY	283	GATCTCAGATCTGTAATGGCTTTTGTGTAAGAGTGTATGGCTTTTCATAGGACTAC	342
Db	359	GAGCTGAAGATCGCCACAGCTCTCGGAGAAATAAGATCTATTTTACAGAAATAT	418
QY	343	ATTGAGTGTGCCAAATAATTACGATGCCAAAGTGGAGCGAGTTGACTTTAGCAATCAT	402
Db	419	TTAGATGCCATCAAGAAATTTTACAGACCACTGCAATCTGTTGATTTTGCAAATGCT	478
QY	403	TTGAGAGACACTAGACGTAATTAATATAGTGGTTGAAATATGAAACACATGCGAAATC	462
Db	479	CCAGAAGAAATCGAAGAAAGATTAACCTCTCGGTGGAAGATCAAAAGAAATGAATAAT	538
QY	463	AAGACGTGATTGTCGAAGTGGCATTAAGCTCATCTGCTGTAATGCTGCTGCTGTAATGCT	522
Db	539	AAATACCTTAATTCCTGAAGTAAATATTGGCAGCAATACACATTTGTTCTTGTGACGCA	598
QY	523	GTGTACTTCAAAGCAAGTGGCAATCAGCTCTCACAGAGCGGAAACCAATAATTCGCAT	582
Db	599	ATCTATTTCAAGGGCAGTGGGAGAGAAATTTAATAAGAGATACATAAAGAGGAAAA	658
QY	583	TTCAAAATCTCCCAAGTCTCTGGGAGGCGAGTCCCATGATGATCATCAGGACCGAATTC	642
Db	659	TTTGTGGCAAAACAGAAATACATACAAGTCCTACAGATGATGAGCAATACATCTTTT	718
QY	643	AAATTTCTCTGTTATCGAGACCCCATCAATCAGATTTCTTCAGCTCAGATACAATGCT---	699
Db	719	CATTTTGGCTTCGCTGGAGATGTACAGGCCAGGTCTCGAAATACCATCAAGGCGAA	778
QY	700	GGCATTAACATGTACGTTCTGCTGCCTGAGA-----ATGACCTCTCTGAATTTGAAAC	753
Db	779	GATCTAAGCATGATTTGTTGTCTGCCAATAAATTCGATGGTCTCCAGAGCTTGAAG	838
QY	754	AAACTGACCTTTCAGAACTCTAATGAATGGACCAATCCAGGCGAATGACCTCTAAGTAT	813
Db	839	AAACTCACTGCTGAGAAATTTGATGAATGGAACAATTTGCAGAAATATGAGAGACACGT	898
QY	814	GTTGAGGTAATTTTTTCTCAGTTCAAGATAGAGAGAAATTAAGAAATGAACAAATATTG	873
Db	899	GTGATTTACATTTACTCTCGTTCAAAGTGAAGAGAGCTATGACCTCAGGACACGTTG	958
QY	874	AGAGCCCTTAGGCTTGAAAGATATCTTTGATGAATCCAAAGCAGATCTCTCTGGGATGCT	933
Db	959	AGAACCATGGGAATGSGTGAATATCTTCAATG---GGGATCGACCTCTCAGGCAATGACC	1015
QY	934	TCGGGGGGTCTGTGTATATATACAGGATGATGACAAATCTTACATAGAGGTCACATGAG	993
Db	1016	GGAGCGCGGTCTCTGCTATCTTGGAGTCTCTACAGAGGCTTTCTTGAGGTTACAGAG	1075
QY	994	GAGGCAACCGAGGTACTGCTGCACAG	1021
Db	1076	GAGGACAGAGAGCTGCAGCTCCACCG	1103

RESULT 14

US-09-266-910-1

```

; Sequence 1, Application US/09266910
; Patent No. 6344362
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Use of a recombinant protein as receptor of a
; TITLE OF INVENTION: hepatitis virus
; NUMBER OF SEQUENCES: 8
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/266,910
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1245 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Squamous Cell Carcinoma Antigen
; TISSUE TYPE: Hepatoma
; CELL TYPE: Hepatocyte
; CELL LINE: HepG2
; US-09-266-910-1

Query Match 18.4%; Score 210; DB 4; Length 1245;
Best Local Similarity 54.4%; Pred. No. 4.8e-58;
Matches 570; Conservative 0; Mismatches 445; Indels 33; Gaps 6;

Qy 1 ATGGCTCCCTCTCTCGACCAATGCAGAGTTTTCCTTCAACCTGTTCCAGAGATGGAT 60
Db 8 ATCAATCACTCGTGAAGCCACACCAAGTTCATGTTCCGACCTGTTCCACAGTTCAG 67

Qy 61 GACAAATCAAGGAATGGAAATGTGTTCTTTTCTCTCTCGAGCTCTTCGTCGCCGTGGCC 120
Db 68 AATCAAAAGAGAA---GAACTCTCTATTCCTCCATCAGCATCACATCAGCATTAGG 124

Qy 121 CTGGTTCGCTTGGGCGCTCAAGTAGCTCCCTCTCTCAGATTGATAAGTTCCTTCAGTT 180
Db 125 ATGGTCCCTCTTAGGAGCCAAAGACACACCTGCACACAGATTAAAGAGGTTCTTCACTTT 184

Qy 181 -----AACACTGCTCAGGATATGGAACCTCTTTAAATGTCAGTCAGG- 224
Db 185 GATCAAGTCACAGAGACACCAAGCTGCAACATATCATGTTGATAGTTCAGGA 244

Qy 225 --GCTCGAGTCTCACTGAAAGAGTTTTTTTCGATATTAATGCATCCCAAGGATTAT 282
Db 245 AATGTTCAATCACCAGTTTCAAAGCTTCTGACTGAATCAACAAATCCCATGATGCATAT 304

Qy 283 GATCTCAGCATGTGAATGGGTTTTTGCTGAAAAGTGTATGGCTTTCTATAAGGACTAC 342
Db 305 GAGCTGAGATCGCGAACAGCTCTTCGAGAAAAAAGCTATCTATTTTACAGGATAT 364

Qy 343 ATTGAGTGTGCGAAAAATTTATACGATGCCAAGTGGAGCGAGTTGACTTTACGATCAT 402
Db 365 TTAGATGCCATCAAGAAATTTTACCAGCAGTGTGGAATCTCTGATTTGCAAAATGCT 424

Qy 403 TTGAAGACACTAGACGTAAATATTAAATAGTGGTTCGAAATGAACACATGGCAAAATC 462
Db 425 CCAGAAGAAGTGAAGAAGATTAACTCTCGGTGGAAAGTCAACGATGAAGAAATTT 484

Qy 463 AAGAACGTCATGTGTAAGTGGCATTAAGCTCATCTGCTTAATGTGTCTGCTGAATGCT 522
Db 485 AANAACCTCTCTCTGAGGTAAATATTGGAGCAATACCAATTTGTTCTTGTGAACGA 544

Qy 523 GTGTATCTTCAAGSCAAGTGGCAATACAGCCCTTCACGAAGAGCGAAACATTAANTGCCAT 582

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 13, 2003, 23:40:42 ; Search time 426 Seconds

(without alignments)
8917.544 Million cell updates/sec

Title: US-09-508-997A-1

Perfect score: 1143

Sequence: 1 atggctcccttgctgcgc.....gcaagtttttgccttga 1143

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2201672 seqs, 1561799599 residues

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1143	100.0	1950	4	US-09-140-719-30
2	1143	100.0	1950	14	US-10-091-442-30
3	399.4	34.9	592	13	US-10-029-386-7445
4	399	34.9	1200	11	US-10-029-386-21145
5	237.8	20.8	1200	11	US-09-823-187-1
6	224.8	19.7	1275	13	US-10-113-113-3
7	220.6	19.3	1279	15	US-10-157-031-354
8	220.6	19.3	1308	9	US-09-735-705-109
9	220.6	19.3	1308	10	US-09-850-716A-109
10	220.6	19.3	1308	10	US-09-897-778-109
11	220.6	19.3	1308	11	US-09-466-396A-109
12	220.6	19.3	1308	13	US-10-117-982-109
13	214.6	18.8	1539	9	US-09-764-898-25
14	211.6	18.5	1711	13	US-10-101-510-669
15	208.8	18.3	1193	8	US-08-731-566-1

16	207.8	18.2	1284	10	US-09-954-456-1841
17	206.8	18.1	1935	15	US-10-106-698-261
18	205	17.9	1419	9	US-09-735-705-111
19	205	17.9	1419	10	US-09-850-716A-111
20	205	17.9	1419	10	US-09-897-778-111
21	205	17.9	1419	11	US-09-466-396A-111
22	205	17.9	1419	13	US-10-117-982-111
23	204.6	17.9	2114	13	US-09-971-392-44
24	204	17.8	1897	15	US-10-208-408-17
25	204	17.8	1908	15	US-10-084-817-27
26	195.4	17.1	1883	10	US-09-895-814-944
27	195.4	17.1	1883	13	US-10-141-678A-944
28	195.4	17.1	1883	13	US-10-294-025-944
29	195.4	17.1	1883	14	US-10-012-896-944
30	194.4	17.0	953	8	US-08-731-566-3
31	193.8	17.0	1549	15	US-10-073-979-12
32	188.4	16.5	1278	13	US-10-094-944-1
33	188.4	16.5	1624	10	US-09-925-300-500
34	187.2	16.4	1465	10	US-09-880-107-3027
35	187.2	16.4	1707	15	US-10-084-817-148
36	186.8	16.3	3603	13	US-10-113-113-1
37	184.4	16.1	1476	10	US-09-974-238-94
38	182.4	16.0	1230	15	US-10-267-311-54
39	182.4	16.0	2847	15	US-10-267-311-20
40	170	14.9	1626	13	US-09-993-363-1
41	158	13.8	1607	9	US-09-729-674-133
42	157.6	13.8	1819	13	US-09-993-363-3
43	156.8	13.7	2566	10	US-09-954-456-1813
44	156.8	13.7	2566	10	US-09-880-107-3283
45	155.2	13.6	1316	10	US-09-969-347-192

ALIGNMENTS

RESULT 1

US-09-140-719-30
; Sequence 30, Application US/09140719
; Patent No. US20010026931A1
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Fuyuki
; APPLICANT: TSURUOKA, No. US20010026931A1DO
; APPLICANT: NAKAZATO, Hiroshi
; APPLICANT: MIURA, Kenju
; APPLICANT: ISHIDA, No. US20010026931A1uniro
; APPLICANT: KURIHARA, Tatsuya
; APPLICANT: YAMAICHI, Kozo
; APPLICANT: YAMAGUCHI, No. US20010026931A1cmi
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/140,719
; FILING DATE: 08-AUG-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/474,661
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993

; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 4-212305
 ; FILING DATE: 17-JUL-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 6-067339
 ; FILING DATE: 04-MAR-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: McGowan, Valcolm K.
 ; REGISTRATION NUMBER: 39,300
 ; REFERENCE/DOCKET NUMBER: 001560-247
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 836-6620
 ; TELEFAX: (703) 836-2021
 ; INFORMATION FOR SEQ ID NO: 30:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1950 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; ORIGINAL SOURCE:
 ; ORGANISM: Homo sapiens
 ; STRAIN: A431
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 1
 ; OTHER INFORMATION: /note= "DNA coding for human
 ; OTHER INFORMATION: megakaryocyte differentiation factor."
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 74..1217
 ; US-09-140-719-30

Query Match 100.0%; Score 1143; DB 9; Length 1950;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTCCCTTGCTGTCAGCAAAATCGAGATTTTGTTCACACCTGTTTCAGAGATGAT 60
 DB 74 ATGGCTCCCTTGCTGTCAGCAAAATCGAGATTTTGTTCACACCTGTTTCAGAGATGAT 133
 QY 61 GACATCAAGGAATGAAATGTGTTCTTCTCTGAGCTCTTCGCTGCCCTGGCC 120
 DB 134 GACATCAAGGAATGAAATGTGTTCTTCTCTGAGCTCTTCGCTGCCCTGGCC 193
 QY 121 CTGGCTCCCTTGCTGTCAGCAAAATCGAGATTTTGTTCACACCTGTTTCAGAGATGAT 180
 DB 194 CTGGCTCCCTTGCTGTCAGCAAAATCGAGATTTTGTTCACACCTGTTTCAGAGATGAT 253
 QY 181 AACACTGCTCAGGATGGAATGAAATGTGTTCTTCTCTGAGCTCTTCGCTGCCCTGGCC 240
 DB 254 AACACTGCTCAGGATGGAATGAAATGTGTTCTTCTCTGAGCTCTTCGCTGCCCTGGCC 313
 QY 241 AAAGAGTTTTTCTGATATAAATGCAATCCCAAGGATGATGATCTCAGCAATGATGAT 300
 DB 314 AAAGAGTTTTTCTGATATAAATGCAATCCCAAGGATGATGATCTCAGCAATGATGAT 373
 QY 301 GGCTTTTCTGATAAAGCTGATGGCTTTCTGATGATGATGATGATGATGATGATGATGAT 360
 DB 374 GGCTTTTCTGATAAAGCTGATGGCTTTCTGATGATGATGATGATGATGATGATGATGAT 433
 QY 361 TTATAGATGTCAGGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
 DB 434 TTATAGATGTCAGGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 493
 QY 421 AATATTAATAGTGGTGTGAAATGAAACACATGCGCAAAATCAAGAACGATGATGATGATGAT 480
 DB 494 AATATTAATAGTGGTGTGAAATGAAACACATGCGCAAAATCAAGAACGATGATGATGATGAT 553
 QY 481 GGTGGCATAGCTCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
 DB 554 GGTGGCATAGCTCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 613

QY 541 TGGCAATCAGCTTTCAGCAGAGCGAAGCAATTAATTCGCAATTTCAATCTCCAGTGC 600
 DB 614 TGGCAATCAGCTTTCAGCAGAGCGAAGCAATTAATTCGCAATTTCAATCTCCAGTGC 673
 QY 601 TCTGGGAAGGCGAGTCGCCATGATGCATCAGGAACGGAAGTTCAATTTGCTGTATTGAG 660
 DB 674 TCTGGGAAGGCGAGTCGCCATGATGCATCAGGAACGGAAGTTCAATTTGCTGTATTGAG 733
 QY 661 GACCCATCAATGAAGATTTCTGAGCTCAGATACAAATGTTGGGCAATAACATGACGTTCTG 720
 DB 734 GACCCATCAATGAAGATTTCTGAGCTCAGATACAAATGTTGGGCAATAACATGACGTTCTG 793
 QY 721 CTGCTGAGATGACCTTCTGAAATTTGAAACAAATGACCTTTCAGAAATCTTAATGAA 780
 DB 794 CTGCTGAGATGACCTTCTGAAATTTGAAACAAATGACCTTTCAGAAATCTTAATGAA 853
 QY 781 TGGACCAATCCAGGCGAATGACCTTTCAGAAATCTTAATGAAATCTTAATGAA 840
 DB 854 TGGACCAATCCAGGCGAATGACCTTTCAGAAATCTTAATGAAATCTTAATGAA 913
 QY 841 ATGAGCAAGATTTATGAATGAAACAAATTTGAGAGCCCTAGGCTGAAAGATATCTTT 900
 DB 914 ATGAGCAAGATTTATGAATGAAACAAATTTGAGAGCCCTAGGCTGAAAGATATCTTT 973
 QY 901 GATGAATCCAAAGCAGATCTCTCGGATTTGCTTCGGGGGCTGCTGTATATATCAAG 960
 DB 974 GATGAATCCAAAGCAGATCTCTCGGATTTGCTTCGGGGGCTGCTGTATATATCAAG 1033
 QY 961 ATGATGCAAAATCTTACATAGAGTCTCAGAGAGGCAACGAGGCTTCTGTCGCA 1020
 DB 1034 ATGATGCAAAATCTTACATAGAGTCTCAGAGAGGCAACGAGGCTTCTGTCGCA 1093
 QY 1021 GGAAGTATATTTAGAAAGCAATCTCCCTCAGTCAGCTCCAGCTGTTAGAGCTGACCA 1080
 DB 1094 GGAAGTATATTTAGAAAGCAATCTCCCTCAGTCAGCTCCAGCTGTTAGAGCTGACCA 1153
 QY 1081 TTCTATTTGTTATCAGAGAGGATGACATCATCTTATTCAGTGGCAAGTTTCTTGCCCT 1140
 DB 1154 TTCTATTTGTTATCAGAGAGGATGACATCATCTTATTCAGTGGCAAGTTTCTTGCCCT 1213
 QY 1141 TGA 1143
 DB 1214 TGA 1216

RESULT 2
 US-10-091-442-30
 ; Sequence 30, Application US/10091442
 ; Publication No. US20020164711A1
 ; GENERAL INFORMATION:
 ; APPLICANT: TSUJIMOTO, Masafumi
 ; INASA, Ruyuki
 ; TSURUOKA, No. US20020164711A1uo
 ; NAKAZATO, Hiroshi
 ; MIURA, Kenju
 ; IGHIDA, No. US20020164711A1uhiro
 ; KURIHARA, Tatsuya
 ; YAMAGUCHI, Kozo
 ; YAMAGUCHI, No. US20020164711A1omi
 ; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
 ; NUMBER OF SEQUENCES: 34
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Burns, Doane, Swecker & Mathis
 ; STREET: P. O. Box 1404
 ; CITY: Alexandria
 ; STATE: Virginia
 ; COUNTRY: United States
 ; ZIP: 22313-1404
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/091,442
 FILING DATE: 07-Mar-2002
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/09/140,719
 FILING DATE: 08-AUG-1998
 APPLICATION NUMBER: US 08/474,661
 FILING DATE: 07-JUN-1995
 APPLICATION NUMBER: US 08/091,028
 FILING DATE: 14-JUL-1993
 APPLICATION NUMBER: JP 4-212305
 FILING DATE: 17-JUL-1992
 APPLICATION NUMBER: JP 6-067339
 FILING DATE: 04-MAR-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: McGowan, Malcolm K.
 REGISTRATION NUMBER: 39,300
 REFERENCE/DOCKET NUMBER: 001560-247
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 836-6620
 TELEFAX: (703) 836-2021
 INFORMATION FOR SEQ ID NO: 30:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1950 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 STRAIN: A431
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 1
 OTHER INFORMATION: /note= "DNA coding for human megakaryocyte differentiation factor."

NAME/KEY: CDS
 LOCATION: 74..1217
 SEQUENCE DESCRIPTION: SEQ ID NO: 30:
 US-10-091-442-30

Query Match 100.0%; Score 1143; DB 14; Length 1950;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGGCTCCCTTCTGTCGAGCAATGCGAGTTTTCCTCAACCTCTTCAGAGATGGAT	60
Db	74	ATGGCTCCCTTCTGTCGAGCAATGCGAGTTTTCCTCAACCTCTTCAGAGATGGAT	133
QY	61	GACATCAAGGAATGGAATGTTCTTTCCTCTCAGATTGATAAGTCTCATGTT	180
Db	134	GACATCAAGGAATGGAATGTTCTTTCCTCTCAGATTGATAAGTCTCATGTT	193
QY	121	CTGCTCCGCTGGGGCTCAAGTCACTCCCTCTCTCAGATTGATAAGTCTCATGTT	180
Db	194	CTGCTCCGCTGGGGCTCAAGTCACTCCCTCTCTCAGATTGATAAGTCTCATGTT	253
QY	181	AACTGCTCAGATATGGAATCTCTTCTTAATAGTCACTCAGGCTCCAGTCTCACTG	240
Db	254	AACTGCTCAGATATGGAATCTCTTCTTAATAGTCACTCAGGCTCCAGTCTCACTG	313
QY	241	AAAGAGTTTTTCTGTATATAATGATCCACAGGATTATGATCAGATTGGAT	300
Db	314	AAAGAGTTTTTCTGTATATAATGATCCACAGGATTATGATCAGATTGGAT	373
QY	301	GGCTTTTGTGAAAAGTGTATGGCTTTTCAAGGACTACATTGAGTGTGCGGAAA	360
Db	374	GGCTTTTGTGAAAAGTGTATGGCTTTTCAAGGACTACATTGAGTGTGCGGAAA	433
QY	361	TTATCAGATGCGCAAGTGGAGGAGTTGACTTTACGAATCAATTAGAGACACTAGAGT	420

Db	434	TTATCAGATGCGCAAGTGGAGGAGTTGACTTTACGAATCAATTAGAGACACTAGAGT	493
QY	421	AATATTAATAGTGGGTTCAAAATGAAACACATGGCAAAATCAAGAACGTGATGGTCAA	480
Db	494	AATATTAATAGTGGGTTCAAAATGAAACACATGGCAAAATCAAGAACGTGATGGTCAA	553
QY	481	GTGCAATAGCTCATCTCTGTATATGGTCTGGTGAATGTGTGTACTTCAAAAGGCAAG	540
Db	554	GTGCAATAGCTCATCTCTGTATATGGTCTGGTGAATGTGTGTACTTCAAAAGGCAAG	613
QY	541	TGGCAATCAGCTTCACCAAGAGGAAACATATAATGGCAATTCAAATCTCCAGTGC	600
Db	614	TGGCAATCAGCTTCACCAAGAGGAAACATATAATGGCAATTCAAATCTCCAGTGC	673
QY	601	TCTGGGAAGCAGTCGCGCATGATGATCAGGAACGGAGTTCAATTTGTCTGTATTTCAG	660
Db	674	TCTGGGAAGCAGTCGCGCATGATGATCAGGAACGGAGTTCAATTTGTCTGTATTTCAG	733
QY	661	GAACCATCAATGAAGATTCTTTGAGCTCAGATACAAATGGTGGCATAAATGATGTTCTG	720
Db	734	GAACCATCAATGAAGATTCTTTGAGCTCAGATACAAATGGTGGCATAAATGATGTTCTG	793
QY	721	CTGCTGAGAAATGACCTCTCTGAAATTTGAAACAACTGACCTTCAGAAATCTAATGAA	780
Db	794	CTGCTGAGAAATGACCTCTCTGAAATTTGAAACAACTGACCTTCAGAAATCTAATGAA	853
QY	781	TGGACCAATCCAAAGCGAATGACCTCTTAAGTATGTTGAGGTATTTTCTCAGTTCAG	840
Db	854	TGGACCAATCCAAAGCGAATGACCTCTTAAGTATGTTGAGGTATTTTCTCAGTTCAG	913
QY	841	ATAGAGAAGAAATTATGAATGAAACAAATATTTGAGAGCCCTTAGGGCTGAAAGATATCTT	900
Db	914	ATAGAGAAGAAATTATGAATGAAACAAATATTTGAGAGCCCTTAGGGCTGAAAGATATCTT	973
QY	901	GATGAATCCAAAGCAGATCTCTCTGGGATTTGCTTCGGGGGGTGTCTGTATATATCAGG	960
Db	974	GATGAATCCAAAGCAGATCTCTCTGGGATTTGCTTCGGGGGGTGTCTGTATATATCAGG	1033
QY	961	ATGATGCAACAATCTTACATAGAGTCACTGAGGAGGGCACCGAGGCTACTGTCGCCACA	1020
Db	1034	ATGATGCAACAATCTTACATAGAGTCACTGAGGAGGGCACCGAGGCTACTGTCGCCACA	1093
QY	1021	GGAGTATATATGTTAGAAAAGCAATCCCTCAGTCCAGCTGTTTATAGCTGACCAACCA	1080
Db	1094	GGAGTATATATGTTAGAAAAGCAATCCCTCAGTCCAGCTGTTTATAGCTGACCAACCA	1153
QY	1081	TTCTATTGTTTATCAGGAAGGATGACATCATCTTATTCAGTGGCAAGTTTCTTGCCCT	1140
Db	1154	TTCTATTGTTTATCAGGAAGGATGACATCATCTTATTCAGTGGCAAGTTTCTTGCCCT	1213
QY	1141	TGA 1143	
Db	1214	TGA 1216	

RESULT 3

US-10-029-386-7445
 ; Sequence 7445, Application US/10029386
 ; Publication No. US20030194704A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Rank, David R.
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; FILE REFERENCE: AEMICA-X-2
 ; CURRENT APPLICATION NUMBER: US/10/029,386
 ; NUMBER OF SEQ ID NOS: 34288
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 7445
 ; LENGTH: 592
 ; TYPE: DNA

;
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/197,525
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/197,087
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1200
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(1191)
US-09-823-187-1

Query Match 20.8%; Score 237.8; DB 11; Length 1200;
Best Local Similarity 52.9%; Pred. No. 6.3e-60;
Matches 608; Conservative 0; Mismatches 502; Indels 39; Gaps 3;

QY 1 ATGGCTCCCTTCTGCGCAAAATGCGAGATTTTGGCTCAACTGTTTCAGAGAGATGGAT 60
DB 7 ATGGACTCTCTTCTTACAGCAACACCAAAATTTTGGCTTGTATCTTTTTCAGAGATGGC 66

QY 61 GACAAATCAAGAAATGGAATGTCTTTTCTCTCTGAGCCCTCTGCTGCCCTGGCC 120
DB 67 AAGATGATCGTCAATAAAACAAATTTTCTCTCCCTGAGCCCTCTGAGCTGCCCTGGT 126

QY 121 CTGGTCCGCTTGGCGCTCAAGATGACTCCCTCTCTGAGATGATGATGATGATGATGAT 180
DB 127 ATGTAGCTTGGTCTGAGAGTGACATGACATCAGATGATGATGATGATGATGATGATGAT 186

QY 181 AACACTGCTCTGAGATATGGAATCTTCTTAATAGTCAATC-----A 222
DB 187 AAGCAATTTTCCCAAGATGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 246

QY 223 GGGCTCAGCTTCACTGAAAGAGTTTCTCTGATATAAATCAATCCCAAGCAAGATTA 282
DB 247 CTGGTCACTGCTACTTTGGCGAGCTTCTCTCCAAATTAGACAGGATCAAGACTGATAC 306

QY 283 GATCTCAGCAATGTAAGTGGCTTTTCTGTAAGAGTGTATGGCTTCTCAAGACTAC 342
DB 307 ACCTGAGTATTGCCACAGGCTTTATGAGAGAGGAATTCCTCAATCTGTGAGGAATAC 366

QY 343 ATTGAGTGTCCGCAAAATATACGATGCAAGTGGAGCGAGTGTGACTTACCAATCAT 402
DB 367 TTAGATGCTGATTCATTTTACCAACAGCAATGAAAGTGTGATTTCCAAATAAC 426

QY 403 TTAGAAGACACTAGACGTAATATTAATAGTGGGTGAAATGAACACATGSCAAATC 462
DB 427 CCTGAAATAATCCAGACAGAGATTAATCTTGGGTGAAATGTCATCCCAAGTAAATC 486

QY 463 AAGAACTGATTTGGTGGAGTGGCAATCACTCTCTGATATGCTGCTGCTGCTGATGCT 522
DB 487 AAGGACTCTTTCAGACAGGACGCTTATATGCTGAGACTGTGCTGCTGCTGCTGCTGCT 546

QY 523 GTGTACTTCAAGGCAAGTGGCAATCAGCTTCCCAAGAGCAACCAATTAATTCGCAAT 582
DB 547 GTTACTTCAAGGCCAATGGGAACATATCTTTCACCAAGAAACACCGCTGGTACCT 606

QY 583 TTCAATCTCCCAAGTGTCTGGGAAGGCAATGCCCAATGATGATCAGAACGGAATTC 642
DB 607 TCTGTCTAAATCAAGATGAAACACAGAGTGTGAAGATGATGACGCAAAAGGCTCTAC 666

QY 643 AATTTGTCTGTTATTGAGGACCCATCAATGAGATTTCTGAGCTCAGATA---CAATGCT 699
DB 667 AGAATTTGCTTATAGAGAGGTGAAGGCAAGATCCTGGAAATGAGTACCAAGGGG 726

QY 700 GGCATAAACATGATGCTTCTGCTGCC-----TGAGATGACCTCTCT 741
DB 727 AAGCTCAGCATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 786

QY 742 GAAATGAAACAAACTGACCTTTTCAGATCTAATGGAATGGACCAATCCAGGGAATG 801

DB 787 GAGCTTGAAGGAAAAATCACCTATGAAAAATGTTGGCTGGAGCAGCTCAGAAAAATG 846
QY 802 ACCCTTAAGTATGTTGAGGTATTTTTTCTCAGTCAAGATAGAGAAATTTATGAATG 861
DB 847 TCAGAAGAAATCGGTGGTCTCTCTTCCCTGCGGTTCACCTGGGAAGACAGCTATGATCTC 906
QY 862 AAACAAATTTTGAAGCCCTAGGCTGAAAGATATCTTTGATGATCCAAAGCAGATCTC 921
DB 907 AATTCCTATTTTACAGACATGGCCATTTACGATATCTTTGATGAAACAGGCGCTGATCT 966
QY 922 TCTGGGATTTCTCGGGGGTCTGCTGATATATATCAAGGATGATGCAATCTTACATA 981
DB 967 ACTGGAATCTCTCAAGTCCCAATTTGATCTTGTCAAAATTTATCCCAAAACCTTTGTG 1026
QY 982 GAGGTCACTGAGGAGGCGCACCGAGCTACTGCTGCCACAGGAAGTAAATTTGTAGAAAG 1041
DB 1027 GAGGTGGATGAAACCGTATCCCGAGGAGCTGACCCACTTGGGCTGTTGCTCGAAAGG 1086
QY 1042 CAACTCCCTCAGTCCAGCTGTTTAGAGCTGACACCCATCTTATTTGTTATCAGGAAG 1101
DB 1087 TCACCTAGATCTTGGGTGGAGTTAATGCCAACACCCCTTTCTCTTTTCATTAGACAC 1146
QY 1102 GATGACATC 1110
DB 1147 AAAAAACC 1155

RESULT 6
US-10-113-113-3
; Sequence 3, Application US/10113113
; Publication No. US20030166852A1
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: PROTEINASE INHIBITOR ZSERP9
; FILE REFERENCE: 00-92
; CURRENT APPLICATION NUMBER: US/10/113,113
; CURRENT FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: 60/280,678
; PRIOR FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1275
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: degenerate sequence
; NAME/KEY: misc feature
; LOCATION: (1)..(1275)
; OTHER INFORMATION: n = A, T, C or G
US-10-113-113-3

Query Match 19.7%; Score 224.8; DB 13; Length 1275;
Best Local Similarity 37.2%; Pred. No. 4.9e-56;
Matches 328; Conservative 154; Mismatches 379; Indels 21; Gaps 2;

QY 248 TTTTCTTGATATAAATCCATCCCAAGGATTAATGATCTCAGCATTTGTAATGGCTTT 307
DB 356 TNYTNWSNARYTNGAYMGNAATHAARACNGAYTACVNYNWSNATHGNCNAYMGNTYT 415

QY 308 TTGCTGAAAGAGTGTATGGCTTTCATAGGACTACATTTGAGTGTGCCGCAAAATATACG 367
DB 416 AYGNHARCARARTYTCNATHGTGCARGARTATYTNAYGNGNATHTCAATTTTAYC 475

QY 368 ATGCCAAAGTGGAGCGGCTGACTTTTACGAATCAATTTAGAGACACTAGAGTAAATTA 427
DB 476 AYACNACNATHGARNSTNGAYTTTCARARAYCCNGHARPNWSNMGNCARGARATHA 535

QY 428 ATAAATGGGTTGAAATGAACACATGCGCAAAATCAAGAACCTGATTTGTTAGAGGTGGCA 487
DB 536 AYTTTGGGTNGARTGYCARWSNCARGGNAARATHAARGARYTNTTYSNARAGYCNRA 595

QY 1106 -ACATCATCTTATTCAGTGGCAGAGTTTCTTCCCTTGA 1143
 |||||
 Db 1168 AACAGCATCTCTTCTTCTGGCAGATTTTCTTCTCTTAA 1206

RESULT 8

US-09-735-705-109
 ; Sequence 109, Application US/09735705
 ; Patent No. US2002005232A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Fengtong
 ; APPLICANT: Fan, Liqun
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: Hosken, Nancy
 ; APPLICANT: Panger, Gary R.
 ; APPLICANT: Li, Samuel X.
 ; APPLICANT: Wang, Aijun
 ; APPLICANT: Skeiky, Yasir A.W.
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: McNeill, Patricia D.
 ; APPLICANT: Panger, Neil

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

; FILE REFERENCE: 210121.455C14

; CURRENT APPLICATION NUMBER: US/09/735,705

; CURRENT FILING DATE: 2000-12-12

; NUMBER OF SEQ ID NOS: 419

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 109

; LENGTH: 1308

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-735-705-109

Query Match 19.3%; Score 220.6; DB 9; Length 1308;
 Best Local Similarity 52.7%; Pred. No. 9e-55;
 Matches 621; Conservative 0; Mismatches 519; Indels 39; Gaps 5;

QY 1 ATGGCTCCCTGCTGCGCAATGAGATTTTGTTCACCTGTTTCAGAGATGGAT 60
 |||||
 Db 25 ATGGATTCACTTGGCGCGTGCAGCTCGACTGGGTGTTGATCTTTCAAGAGCTGAG 84
 |||||
 QY 61 GACATCAAGGAATGGAATGTCTTTCTCTCTCTCAGATGATAAGTTGCTTATGTT 180
 |||||
 Db 142 ATGGTCTCTCTGGGACCCGAGGACCCAGCTTCCAGTGGAGGAGTGTTCACCTCT 201
 |||||
 QY 181 AACCTGCTCA-----GGATGGAACCTTCTTAATAGTAG 219
 |||||
 Db 202 GAAAAGAGCAAGAGCTCAAGATAAAGGCTGAAGAAAAGAGGTGATTGAGAACA 261
 |||||
 QY 220 TCAGGCTCCAGTCTCACTGAAAAGAGTCTTTCTGATATAATGATGATCCCAAGAT 279
 |||||
 Db 262 GAGCAGTACATCAATCCAAAGTTTTCAGCTGAAATGAAACAACTCAATATGAT 321
 |||||
 QY 280 TATGATCTCAGATTTGATAGGCTTTTGTCTGAAAAGTGTATGCTTTGATAGGAC 339
 |||||
 Db 322 TATGAATGAACATAACCAACAGGCTGTTGGAGAAAACATACCTCTCTTCTTCAAAA 361
 |||||
 QY 340 TACATTAGTGTCCAAAATATATAGATGCCAAAGTGGAGCGAGTGTACTTACGAAT 399
 |||||
 Db 382 TACTTAGATTATGTGAAAATATATATCATGCTCTCTGAAACCTGTGATTTGTAAT 441
 |||||
 QY 400 CATTTAGAGACACTAGAGTATATTAATAGTGGTTGAAATGAAACATGCGAAA 459
 |||||
 Db 442 GCAGCCGATGAGTGTGAAGAGATTAATTCCTGGGTGAAACAAACAAATGAAAA 501
 |||||
 QY 460 ATCAAGAACGTGATTGGTGAAGGTGGCATAAGCTCATCTGCTGTATGTTGCTGAT 519

Db 502 ATCAAGCACTTGTCCAGATGGCTCTATTAGTAGCTTACCAAGCTGCTGTGTGTAAC 561
 |||||
 QY 520 GCTGTGTACTTCAAGGCAAGTGGCAATCAGCTTCCACAGAGCGGAACATAATGTC 579
 |||||
 Db 562 ATGGTTTATTTAAAGGCAATGGGACAGGGAGTTTAAAGAAAGAAATATCTAAGGAAG 621
 |||||
 QY 580 CATTTCAAAATCTCCAAAGTGTCTGGGAAAGGAGTGGCCATGATCAGGAACGGAAG 639
 |||||
 Db 622 AATTTTGGATGAATAAGAGCACAGTAATCTGTACAGATGATGACACAGGCAATTC 681
 |||||
 QY 640 TTCAATTTGTCTGTTATTTAGAGCCCATCAATGAAGATTTCTGAGCTCAGAT---ACAAT 696
 |||||
 Db 682 TTTAGCTTCACTTTCTCTGGAGACCTTGCAGGCCAAAATTTAGGGATTCATATAAAAAC 741
 |||||
 QY 697 GGTGGCATAAAACATGATCTGCTGCTGCTGTA-----GAATGACCTCTTGAATTTGAA 750
 |||||
 Db 742 AACGACCTAAGCATGTTTGTGCTTCTGCCCAACGACATCGATGGCTGGAGAGATAATA 801
 |||||
 QY 751 AACAACTGACCTTTTCAAGATCTAATGAAATGGACCAATCCAAAGGCAATGACCTCTAAG 810
 |||||
 Db 802 GATAAATAAGTCTCTGAGAAATTTGTAGATGGACTAGTCCAGGSCATATGGAAGAAAGA 861
 |||||
 QY 811 TATGTTCAAGTATTTTCTCTCAGTTCAGATGAGAGAAATATGAAATGAACAAATAT 870
 |||||
 Db 862 AAGTGAATCTGCATTTGCCCTGGGTTTGGAGTGGAGGACAGTTTACGATCTAGAGCGGTC 921
 |||||
 QY 871 TTGAGAGCCCTAGGCTGAAAGATATCTTTGATGAATCCAAAGCAGATCTCTCTGGGATTT 930
 |||||
 Db 922 CTGCTGCCATGGGATGGGCGATGCTTCACTAGGACAAAGCCCACTACTCTGGGATG 981
 |||||
 QY 931 GCTTCGGGGGCTGCTGTATATATCAAGATGATGACAAATCTTTACATAGAGTCACT 990
 |||||
 Db 982 TCGTCAAGCTCCGGTTGTACGCCAGAAATTCCTGCACAGTTCTTTTGTGCGAGTAACT 1041
 |||||
 QY 991 GAGGAGGCAACCGAGGCTACTGCTCCACAGAAATGATATGTAAGAAAGCAATCTCCCT 1050
 |||||
 Db 1042 GAGGAGGCAACCGAGGCTGAGCTGCCACTGGCATAGGCTTTACTGTCACTCCGCCCA 1101
 |||||
 QY 1051 CAGTCCACGCTGTTTAGAGCTGACCAACCACTCTCTTTGTTATCAGGAAGT----- 1105
 |||||
 Db 1102 GGTCATGAATGTTTCACTCAATCATCTCTCTCTCTCTCTCTCATCAGGCACATGAATC 1161
 |||||
 QY 1106 -ACATCATCTTATTCAGTGGCAAGTTTCTTGCCTTGA 1143
 |||||
 Db 1162 AACAGCATCTCTTCTTGGCAGATTTTCTTCTCTCTTAA 1200

RESULT 9

US-09-850-716A-109

; Sequence 109, Application US/09850716A

; Patent No. US20020115139A1

; GENERAL INFORMATION:

; APPLICANT: Kalos, Michael D.

; APPLICANT: McNeill, Patricia D.

; APPLICANT: Retter, Marc W.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

; FILE REFERENCE: 210121.455C15

; CURRENT APPLICATION NUMBER: US/09/850,716A

; CURRENT FILING DATE: 2001-05-07

; NUMBER OF SEQ ID NOS: 44C

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 109

; LENGTH: 1308

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-850-716A-109

Query Match 19.3%; Score 220.6; DB 10; Length 1308;
 Best Local Similarity 52.7%; Pred. No. 9e-55;
 Matches 621; Conservative 0; Mismatches 519; Indels 39; Gaps 5;


```
Db 502 ATCAAGCACTGTTCCCAAGATGCGCTCTATTAGTAGCTCTACCAAGCTGGTGTGTAAC 561
Qy 520 GCTGTGTAATCAAAAGGCAAGTGGCAATCAGCTTTCACAGAGCCAAACCTAAATTCG 579
Db 562 ATGCTTTATTTTAAAGGCAATGGACAGGAGTTTAAAGAAAGAAATCTAAGAAAG 621
Qy 580 CATTTTAAATTCCTCAAGTGTCTGGGAAGGAGTGGCCATGATGCAATCAGGAACGGAAG 639
Db 622 AAATTTTGGATGAATAAGACACAAAGTAATCTGTACAGATGATGACACAGGCATTC 681
Qy 640 TTCAATTTCTGTGTTATGAGAGCCCAATCAATGAGAGATTTTGGAGTCTAGAT 696
Db 682 TTATGCTTCACTTTCTCGAGGACTTGCAGGCCAAATTTAGGGATTCATATAAAAC 741
Qy 697 GGTGGCAAAAATCTAGCTTCTGCTGCTGA-----GAATGAOCTCTCTGAAATGAA 750
Db 742 AACGACTTAAGCAATTTTGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 801
Qy 751 AACAACTGACCTTTTCAAGATCTAATGGAATGGACCAATCCAAAGGCAATGACCTTAAG 810
Db 802 GATAAATAAAGTCTGCAAAATTTGGTAGAGTGAATAGTCCAGGGCATATGAGAAAGA 861
Qy 811 TAGTGTGAGTATTTTCTCTCAGTTTCAGATAGAGAGAAATTTGAATGAACAAATAT 870
Db 862 AAGTGGAATCTGCACTTGGCCCGGTTTGGAGTGGAGGACAGTTTACGATCTAGAGCGGTC 921
Qy 871 TTGAGAGCCCTAGGCTGAAAGATATCTTTGATGAATCCAAAGCAGTCTCTCTGGGATT 930
Db 922 CTGGCTGCCATCGGAGTGGGATGCTTTCAGTGAACCAAGGATTCCTTTGGGCAATCT 1041
Qy 931 GCTTGGGGGTCGTGTGTATATATCAAGGATGATGCAAAATCTTACATAGAGTCACT 990
Db 982 TCGTCAAGCTCTCGGCTTGTACGCCAGAAATCTCTGACAGTTCTTTTGGGCAATCT 1041
Qy 991 GAGGAGGCAACGAGCTACTGCTGCCACAGGAAGTAATTTGTAGAAAGCACTCCCT 1050
Db 1042 GAGGAGGCAACGAGGCTGAGCTGCCACAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1101
Qy 1051 CAGTCCACGCTGTTTAGAGCTGACCAACCAATCTTATTTGTTATCAGGAGGATG---- 1105
Db 1102 GGTATGAAATGTTTCACTGCAATCATCCCTTCTGTTCTTTCATGAGCAGATGATCC 1161
Qy 1106 -ACATCACTTATTCAGTGGCAAGTTCCTTGGCCCTTGA 1143
Db 1162 AACAGCATCTCTTCTCGGCAGATTTCTTCTCCTTAA 1200

RESULT 11
US-09-466-396A-109
; Sequence 109, Application US/09466396A
; Publication No. US20030119763A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C4
; CURRENT APPLICATION NUMBER: US/09/466.396A
; CURRENT FILING DATE: 1999-12-17
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 109
; LENGTH: 1308
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-466-396A-109

Query Match 19.3%; Score 220.6; DB 11; Length 1308;
Best Local Similarity 52.7%; Pred. No. 9e-55;
Matches 621; Conservative 0; Mismatches 519; Indels 39; Gaps 5;

Qy 1 ATGGCTCCCTTGTGTCAGCAATCCAGAGTTTGTCTCAACCTGTTTCAGAGATGAT 60
Db 25 ATGGATTCACCTGGCCGCTGAGCACTCGACTGGGTTTGATCTTTTCAGAGCTGAAG 84
```

```
Qy 61 GACATCAAGAAATGGAATGTCTTTTCTCTCTGAGCTCTCTGCTCCCTGCC 120
Db 85 AAAACAAATG---ATGCAACATCTTCTTTTCCCTGTGGGCATCTTGACTGCAATGGC 141
Qy 121 CTGCTCCGCTTGGCGCTCAAGATGACTCCCTCTCTCAGATTTATAGTTGCTTCATGTT 180
Db 142 ATGGTCTCTCTCTGGGACCCGAGGAGCCACCGCTTCCAGTTTGGAGAGGTGTTCACTCT 201
Qy 181 AACACTGCTCA-----GGATATGAAACTCTTCTTAATAGTCAG 219
Db 202 GAAAAGAGAGAGAGAGCTCAAGATTAAGAGCTGAAGAAAAGAGGTGATTAAGAACACA 261
Qy 220 TCAGGGCTCAGCTCTCACTGAAAGAGTTTTTTTCTGATATAAATGATCCACAGGAT 279
Db 262 GAGCAGTACATCAAAATTTCCAAAAGTTTTTGAATTAAGCAAACTCAATATGAT 321
Qy 280 TATGATCTCAGCATTTGTAATGGGCTTTTGTGAAAAGTGTATGGCTTCTATPAGGAC 339
Db 322 TATGATCTGACATTAACCAAGGCTTTTGGAGAAAATAATACCTCTCTCTTCAAAA 381
Qy 340 TACATTCAGTGTGCCAAAAATTTATAGATGCAAAAGTGGAGGAGTTCGACTTTACGAAT 399
Db 382 TACTTAGATTTGTTGAAAAATTTATCAAGCATCTCTGGAACCTCTGTGTTTGTAAAT 441
Qy 400 CATTTAGAAGACACTAGACGTATATTTAATAGTGGTTGAAAATGAACACATGGCAA 459
Db 442 GCACCCCATGAAGTCGAAAGAGATTAATCTCGGTTGAAAGCAAAACAAATGAAAA 501
Qy 460 ATCAAGAACTGATTTGGTGAAGTGGCATTAAGTCACTCTCTCTTANTGTGCTGCAAT 519
Db 502 ATCAAGCACTTCTCCAGATGCTCTATTAGTACTCTTACCAGCTGTCTGTCTGTGTAAC 561
Qy 520 GCTGTGACTTCAAGGCAAGTGGCAATCAGCTTTCACCAAGAGCAAAACCAATATTCG 579
Db 562 ATGCTTTTAAAGGCAATGGACAGGAGTTTAAAGAAAGAAATTAATAAGAGAG 621
Qy 580 CATTTCAATCTCCCAAGTCTCTGGAGAGGAGTGGCCATGATGATCAGGAACGGAAG 639
Db 622 AAATTTGGATGAATAGAGACAAAGTAATCTGTACAGATGATGACACAGAGCATTCC 681
Qy 640 TTCAATTTGCTGTTATGAGGACCCATCAATGAGAGATTTCTGAGCTCAGAT---ACAT 696
Db 682 TTTAGCTTCACTTCTCTGGAGACTTGGAGGCCAAATTTTAGGGATTCATATAAAAC 741
Qy 697 GGTGGCATAAACATGACTGTTCTGCTGCTGA-----GAATGACCTCTCTGAAATGAA 750
Db 742 AACGACTTAAGCATGTTTGTGCTTCTGCCAACGACATCGATGGCTGGAGAGATAATA 801
Qy 751 AACAACTGACCTTTCAGATCTAATGGAATGGACCAATCCAAAGGCAATGACCTCTAAG 810
Db 802 GATAAATAAAGTCTGAGAAATTTGGTAGAGTGAATAGTCCAGGGCATATGAGAAAGA 861
Qy 811 TATGTTGAGGTATTTTTTCTCTCAGTTTCAGATAGAGAAATTTATGAAATGAAACATAT 870
Db 862 AAGTGAATCTGCACTTGGCCCGGTTTGGAGTGGAGCAAGTTTACATCTAGAGCGGTC 921
Qy 871 TTGAGAGCCCTAGGCTGAAAGATATCTTTGATGAATCCAAAGCAGATCTCTCTGGGAT 930
Db 922 CTGGCTGCCATGGGATGGCGATGCTTTCAGTGAACAAAGCCGACTACTCTGGGAATG 981
Qy 931 GCTTCCGGGGTCTGCTGTATATATCAAGGATGATGCAAAATCTTACATAGAGTCACT 990
Db 982 TCGTCAAGCTCTCGGTTTGTACGCCAGAAATTTCTGCAAGTTCTTCTTGGCAGTAACT 1041
Qy 991 GAGGAGGCAACGAGGCTACTGCTGCCACAGGAAGTAATTTGTAGAAAGCACTCCCT 1050
Db 1042 GAGGAGGCAACGAGGCTGAGCTGCCACAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1101
Qy 1051 CAGTCCACGCTGTTTAGAGCTGACCAACCAATCTTATTTGTTATCAGGAGGATG---- 1105
Db 1102 GGTATGAAATGTTTCACTGCAATCATCCCTTCTGTTCTTTCATGAGCAGATGATCC 1161
Qy 1106 -ACATCACTTATTCAGTGGCAAGTTCCTTGGCCCTTGA 1143
Db 1162 AACAGCATCTCTTCTCGGCAGATTTCTTCTCCTTAA 1200
```

DD ATCAAGGACATGGTTCACAGATGGCTCTATATAGTAGCTCTACCAAGCTGGTGCTGGTGAAC 561

QY 310 GCTGAAAGTGTATGCTTTTCATAGAGCTACATTGAGTGTGCGCAAAATATACGAT 369
Db 177 GGGACAAAGACGATGGCATTTTCATAGCAATATTTAAGCTGTTCTGAGAAATGGTATCAA 236
QY 370 GCCAAGTGGAGCGAGTTGACATTTACGAATCATTTAGAAGACACTAGACGTATATTAAT 429
Db 237 GCCAGTTTGCAACTGTGGATTTTGAACAGCTCTACAGAGAAACGAGGAAACGATTAT 296
QY 430 AAGTGGTGTAAATGAAACATATGCGAAATCAAGACGTCATTTGTGAGTGGCAT 489
Db 297 GCTTGGTGTAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 356
QY 490 AGTCATCTGCTGTATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 549
Db 357 GACCTTTATCTGTATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 416
QY 550 GCTTTCACCAAGAGCGAAACCAATTTGCAATTTCAATTTCAATTTCAATTTCAATTT 609
Db 417 AAATTTCAAGTAAAGAGACAGTATAAAGTCTTTTACGCTAAGTGAAGTAAATGTA 476
QY 610 GCAGTGGCATGATGATCAGAACGAGAGTCAATTTGCTGTGCTGTGCTGTGCTGTGCT 669
Db 477 ACTGTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 536
QY 670 ATGAAGATCTTTCAGCTCAGATACATATGCTGCTAATGCTGCTGCTGCTGCTGCTG 729
Db 537 ATGCAAGTCTTTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 596
QY 730 AATGACCTCTCTG-----AAATGAAACAAACCTGACCTTTTCAGAACTCTTAATGGA 780
Db 597 GTAGGCATAGCTAATCTGAAACAGATAGAAAGCAGCTGAATTCGGGAGCTTTCATGAG 656
QY 781 TGGACCAATCAAGGCGATGACCTTAAGTATGCTGAGTATTTTTCCTCAGTTCAG 840
Db 657 TGGACCAAGCTCTTCTACATGATGAAAGAGAGTGAAGTACACCTCCCGAGTTCAA 716
QY 841 ATAGAGAAAGATTTAGAAATGAAACATATTTGAGAGCCTGAGGCTGAAAGATCTTT 900
Db 717 CTGAAATTAAGTATGCTAATTTCCCTGTTAAACCTCCAGGGGTGACAGATCTCTTC 776
QY 901 GATGAATCAAAGCAGATCTCTCTGGATTTGCTTGGGGGGTCTGTGTATATATCAAG 960
Db 777 AACCAAGTCAAAGCTGATCTTTTGGAAATGTCACCAACCAAGGCTCTATTTATCAA 836
QY 961 ATGATGCAAAATTTACATAGAGTCTGAGAGGGGACCGAGGCTCTGCTGCGACA 1020
Db 837 GCATCCCAAGTCTACCTGATGTACGGAAGAGGGCAGGAGGAGCAGCAGCCACT 896
QY 1021 GGAATTAATTTAGAAAGCAACTCCCTCAGTCCACGCTGTTTAGAGTGCACCA 1080
Db 897 GGGACAGCATCGCTGTAAAGAGCTTACCAATGAGAGCTCAGTTCAAGGGCAACACCC 956
QY 1081 TTCCATTTTGTATCAGGAAGGATGATCATC 1113
Db 957 TTCCTGTTCTTTTAAAGGCACACTCATACCAAC 989

RESULT 14

US-101-510-669

; Sequence 669, Application US/10101510

; Publication No. US20030148295A1

; GENERAL INFORMATION:

; APPLICANT: WAN, JACKSON

; APPLICANT: WANG, YIXIN

; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE

; FILE REFERENCE: 15117.0012

; CURRENT APPLICATION NUMBER: US/101/101,510

; CURRENT FILING DATE: 2002-03-20

; PRIOR APPLICATION NUMBER: 60/276,947

; PRIOR FILING DATE: 2001-03-20

; NUMBER OF SEQ ID NOS: 805

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 669

; LENGTH: 1711
; TYPE: DNA
; ORGANISM: Homo sapiens
US-101-510-669

Query Match 18.5%; Score 211.6; DB 13; Length 1711;
Best Local Similarity 54.5%; Pred. No. 5.2e-52;
Matches 571; Conservative 0; Mismatches 444; Indels 33; Gaps 6;

QY 1 ATGCGCTCCCTGCTGCGCAAAATGACAGATTTTGTCTTCAACTGCTTCAGAGAGATGGAT 60
Db 62 ATGANTTCTACTAGTGAAGCCACACCAAGTTCATGTTGACGCTGTTCCACAGTTCAGA 121
QY 61 GACAAACAAAGGAATGGAATGTTTTCCTCTCTGAGGCTCTTCGTCGCCCTGGCC 120
Db 122 AAATCAAAAGAAA---CAACATCTTCTATTCCTCCCTATCAGCATCACATCAGCATTAGGG 178
QY 121 CTGGTCCGCTTGGGCGCTCAGATGATCTCTCTCAGATTGATAGTTGCTTCAATGTT 180
Db 179 ATGGTCTCTTAGGAGCCAAAGACAACTGCAACACAGATTAAAGAGGTTCTTCACTTT 238
QY 181 -----AACACTGCTCAGGATATGAAACTCTTCTTAAATAGTCAGTCAGG- 224
Db 239 GATCAAGTCACAGAGACACACAGGAAAGCTGCAACATATCATGTTGATAGTCAGGA 298
QY 225 --GCTCCAGTCTCAACTGAAGAAGTTCCTGATATAAATGCAATCCCAAGGATTAT 282
Db 299 AATGTTTCATCACCAGTTTCAAAGCTTCTGACTGAATTCACAAATCCACTGATGATAT 358
QY 283 GATCTCAGCATTTGGAATGGCTTTTCTGTAAGAGTGTATGGCTTTTCATAGGACTAC 342
Db 359 GACTGAGATCGGCAACAGCTCTTCGAGAGAAACAAACGATATCTATTTTACAGGAATAT 418
QY 343 ATTGACTGTGCGGAAAAATATACGATCCAAAGTGGAGCGAGTTGACTTTTACGAATCT 402
Db 419 TTAGTGGCATCAAGAAATTTACGACACAGTGTGAATCTGTTGATTTGCAATGCT 478
QY 403 TTAGAGACACTAGACCTATATTAATGAGTGGTGTAAATGAAACACATGCGCAAAATC 462
Db 479 CCAGAGAAAGTCGAAAGAGATTAACCTCTGGGTGAAAGTCAACGAATGAAAAAT 538
QY 463 AAGACGCTGATTGTGAGGTGCGCATAGCTCATCTCTGTTAATGGTCTGTGCTGCTGCT 522
Db 539 AAAACCTTATTTCTGAGGTATATTTGGCAGCAATACCACTGTTGTTGTAAGCA 598
QY 523 GTGTACTTCAAAGGCAAGTGGCAATCAGCTTTCACAGAGCGAAACCAATAATTTGCCAT 582
Db 599 ATCTATTTCAAAGGGCAGTGGGAGAGAAATTTAATAAAGAGATCTATAAGAGGAAAA 658
QY 583 TTCAATCTCCCAAGTCTCTGGAGGAGCTGCGCATGATGATGATGATGATGATGATGAT 642
Db 659 TTTTGCCAAAACAAATATACATCAAGTCCATACAGATGATGAGGCAATACATCTTTT 718
QY 643 AATTTCTCTGTTTATGAGGACCCATCAATGAGATTTCTGAGCTCAGATCAGATGTT--- 699
Db 719 CATTTTGGCTCGCTGGAGAGTGTACAGCCAGGTCTGGAATACCAATCAAGGCAAA 778
QY 700 GGCATTAACATGTACGTTCTGCTGCTGAG- ----ATGACCTCTCTGAAATTTGAAAC 753
Db 779 GATCTAAGCATGATTGTGTGCTGCCAAATGAAATCGATGGTCTCCAGAGCTTGAAGAG 838
QY 754 AAATGACCTTTCAAGATCTAATGGAATGAGCAATCCAGGCGCAATGACCTCTAAGTAT 813
Db 839 AAATCACTGCTGAGAAATGATGGAATGGAAGTTCGCAATATGAGAGGACACGT 898
QY 814 GTTGAAGTATTTTCTCAGTTCAAGATGAGAGAAATTTATGAAATGAAACATATTG 873
Db 899 GTCGATTACCTTACCTCGTTTCAAGTGAAGAGAGCTATGACCTTCAAGGACAGTTG 958
QY 874 AGAGCCCTTAGGCTGAAAGATATCTTTGATGAATCCAAAGCAGATCTCTCTGGGATGCT 933
Db 959 AGAACCCTGGAAATGGTGGATATCTTCAATG---GGGATGACAGCTCTCAGGCAATGACC 1015

QY 934 TCGGGGGTCTCTTATATCAAGGATGATCGAATAATCTTACATAGAGTCACTGAG 993
 Db 1016 GGGAGCCGGTCTCGTGTATCTCGAGTCTACACAAGGCTTTGTGGAGTTACAGAG 1075
 QY 994 GAGGCGACCGAGCTACTGCTGCCACAG 1021
 Db 1076 GAGGAGCAGAGCTGCGAGTCCACCG 1103

RESULT 15

US-08-731-566-1
 ; Sequence 1, Application US/08731566
 ; Publication No. US20030072752A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Toshiko Hibino et al.
 ; TITLE OF INVENTION: TWO NOVEL GENES FROM PSORIATIC EPIDERMIS:
 ; TITLE OF INVENTION: PSORIASTATIN TYPE I AND PSORIASTATIN TYPE II
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD, LLP
 ; STREET: 28 State Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02109-1775
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/731,566
 ; FILING DATE: 16-OCT-1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/013,755
 ; FILING DATE: 20-MAR-1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/005,679
 ; FILING DATE: 17-OCT-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Louis Myers
 ; REGISTRATION NUMBER: 35,965
 ; REFERENCE/DOCKET NUMBER: MGP-040CP
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617)227-7400
 ; TELEFAX: (617)227-5941
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1193 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 22..1193
 ; US-08-731-566-1

Query Match 18.3%; Score 208.8; DB 8; Length 1193;
 Best Local Similarity 53.7%; Pred. No. 2.8e-51;
 Matches 563; Conservative 0; Mismatches 452; Indels 33; Gaps 5;

QY 1 ATGGCTCCCTTGGTCGAGCAATGACAGATTTTGGTTCAACTGTTTCAGAGAGATGGAT 60
 Db 22 ATGAATCTCAGTGAAGCAACACCAAGTTTCATGTTGCACTGTTCCACAGTTTCA 81
 QY 61 GACAATCAAGGAATGAAATGTGTTCTTTCTCTGAGGCTCTTGGTGCCTGGGC 120
 Db 82 AAATCAAAAGAGAA--CAACAATCTTATTCCTATCAGCATCAGTATAGGG 138
 QY 121 CTGTCGGCTGGGGCTCAAGATGACTCCCTCTCTCAGATGCAATGCAATGCTCTCATGTT 160
 Db 139 ATGTCCTCTTAGGAGCCAAAGGCAACACTGCACAACAGATTAAAGAGGTTCTTCACTTT 198

QY 181 AACATGCTCAGGATATGGA-----ACTTCTTATATAGTCACTCA 222
 Db 199 GATCAAGTTCACAGAGAACACCAAGAAAAGCTGCAACATATCATGTGTATAGTCAAGA 258
 QY 223 GGGCTCCAGTCTCAACTGAAAGAGTGTCTTCTGATATAAATGCATCCACAGGATAT 282
 Db 259 GATGTTATCATCAGAGTTTCAAGAGCTTCTGATGATTCACAAATCACTGATGATAT 318
 QY 283 GATCTCAGCATTTGAAATGGGCTTTTCTGAAAAGTGTATGGCTTTTATAAGGCTAC 342
 Db 319 GAGCTGAGATCGCCAAACAGCTCTCGGAGAAAACGATATCTATTTTACAGGAATAT 378
 QY 343 ATTGAGTGTCCGAAAATTTATACAGTGCCTGAGAGGAGGAGTGTGACTTACGATCAT 402
 Db 379 TTGATGTCATCAAGAAATTTTACAGACCAAGTGTGGAATCTGTTGATTTTCAGATGCT 438
 QY 403 TTAGAAGACACTAGACGCTAATTAATTAAGTGGGTGAAAATGAAACACATGCGAAATC 462
 Db 439 CCAGAGAAAGTCCGAAAGAGATTAATCTCTGGGTGAAAGTCAACGAAATGAAATAT 498
 QY 463 AAGAACGTGATTTGGAAGTGGCATTAAGCTCATCTCTGCTGTAATGCTGCTGCTGATGCT 522
 Db 499 AAAAACTTAATTCCTGAAGGTAATATGGCAGCAATACCAATGCTGTTCTTGTGAAGCA 558
 QY 523 GTGTACTTCARAGGCAAGTGCATCAGCTTCCACAGAGGCAACCATAAATTTGCCAT 582
 Db 559 ATCTATTTCAAAGGGCAGTGGGAGAGAAATTTAATAAGAGATACATAAGAGGAAAA 618
 QY 583 TTCAAAATCTCCCAAGTGTCTCTGGAAGGAGTGCCTGATGATGATGATGATGATGATGAT 642
 Db 619 TTTTGCCCAACACAGAAATACATACAGTCCATACAGATGATGAGGCAATACACATCTTT 678
 QY 643 AATTGTCTGTTTATGAGGACCCATCATAGAGATTTCTTGAGCTCAGATACAAATGCT--- 699
 Db 679 CATTTTGGCTCTGCTGAGGATGTACAGGCCAAGGTCTGGAATACCATACAAAGGCAAA 738
 QY 700 GGCATAAACATGTACGTTCTGCTGCTGAGA-----ATGACCTCTCTGAAATGAAAC 753
 Db 739 GATCTAAGCAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 798
 QY 754 AAACAGACCTTTTCAAGAAATCTAATGGAATGGACCAATCCAAAGGCAATGACCTCTAAGTAT 813
 Db 759 GAACCTCACTGCTGAGAAATGATGGAATGGACCAAGTTTGCAGAAATATGAGAGAGACACGT 858
 QY 814 GTTGAGGTATTTTCTCTCAGTTCAAGATAGAGAGAAATTTATGAATGAAACAAATTTTG 873
 Db 859 CTCGATTTTACCTTACCTCGGTTCAAAGTGGAGAGAGAGCTATGACCTCAAGGACACGTTG 918
 QY 874 AGAGCCTAGGGCTGAAAGATATCTTTGATGAATCCAAAGCAGATCTCTCTGGGATGCT 933
 Db 919 AGAATCATGGAAATGTTGATATCTTCAATG---GGGATCAGACCTCTCAGGCAATGACC 975
 QY 934 TCGGGGGTCTCTGTATATATCAAGATGATGCAAAATCTTATAGAGGTCACTGAG 993
 Db 976 GGGAGCCGGTCTCTGCTATCTGAGTCTGAGGCTCTGCAAGGCTTTTGTGGAGTTACAGAG 1035
 QY 994 GAGGCGACCGAGCTACTGCTGCCACAG 1021
 Db 1036 GAGGAGCAGAGCTGCGAGTCCACCG 1063

Search completed: December 14, 2003, 01:13:19
 Job time : 437 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2003, 00:59:44 ; Search time 2773 Seconds
(without alignments)
10018.037 Million cell updates/sec

Title: US-09-508-997A-1

Perfect score: 1143

Sequence: 1 atggctcccttgctgcgc.....gcaagtttttgccttga 1143

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hrc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hrc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	781.4	68.4	1889 11	AK014524 Mus muscu
2	747.8	65.4	786 10	SG185954 RST4909 A
C 3	604.4	52.9	638 9	AW381394 PMO-HT030
C 4	601	52.6	638 9	AW381390 PMO-HT030

5	559.4	48.9	1026	10	BG682538	BG682538 602624811
6	536	46.9	824	12	BI871335	BI871335 60332761
7	497.2	43.5	945	14	BY713938	BY713938 BX713938
8	496.4	43.4	498	13	EX284040	EX284040 EX284040
9	451.2	39.5	906	14	CB590890	CB590890 AGENCOURT
10	405	35.4	405	10	BG319597	BG319597 4chr19 Hum
11	384.4	33.6	527	10	BF477923	BF477923 61859005
12	321.8	28.2	363	14	D82606	D82606 HUMHBC2926
13	321.8	28.2	363	14	T48573	T48573 hbc2926 Hum
14	262.8	23.0	428	14	CB794396	CB794396 AMGNNUC:S
15	234.2	20.5	1688	11	AK009003	AK009003 Mus muscu
16	234.2	20.5	1700	11	AK009855	AK009855 Mus muscu
17	213.2	18.7	1619	11	AK040697	AK040697 Mus muscu
18	213.2	18.7	1688	11	AK009018	AK009018 Mus muscu
19	211.6	18.5	1736	11	BC008483	BC008483 Homo sapi
20	209.6	18.3	1609	11	AK030670	AK030670 Mus muscu
21	208.8	18.3	1422	11	AK084785	AK084785 Mus muscu
22	207.2	18.1	1344	11	BC008399	BC008399 Homo sapi
23	204	17.8	1633	11	AK003220	AK003220 Mus muscu
24	196.6	17.2	1962	11	AK078753	AK078753 Mus muscu
25	196	17.1	381	13	BY139401	BY139401 BY139401
26	187.2	16.4	896	10	EG682440	EG682440 602630184
27	186	16.3	186	12	BI335844	BI335844 8d119 Hum
28	181	15.8	1289	11	AK003930	AK003930 Mus muscu
29	178.6	15.6	1662	11	AK003650	AK003650 Mus muscu
30	178.4	15.6	1289	11	AK008914	AK008914 Mus muscu
31	178.4	15.6	3468	11	AK036537	AK036537 Mus muscu
32	178	15.6	2940	11	AK036811	AK036811 Mus muscu
33	177.8	15.6	1289	11	AK076129	AK076129 Mus muscu
34	175.2	15.3	1719	11	BC011076	BC011076 Mus muscu
35	173.2	15.2	754	13	EX099033	EX099033 EX099033
36	173	15.1	1266	11	AK052580	AK052580 Mus muscu
37	168.2	14.7	2102	11	AK028382	AK028382 Mus muscu
38	164.4	14.4	1952	11	AK090049	AK090049 Mus muscu
39	164.4	14.4	1965	11	AK081487	AK081487 Mus muscu
40	161.8	14.2	1693	11	AK018226	AK018226 Mus muscu
41	161.4	14.1	1922	11	AK049429	AK049429 Mus muscu
42	160.4	14.0	2671	11	AK032659	AK032659 Mus muscu
43	159.6	14.0	858	10	BG199939	BG199939 RST19235
44	159.2	13.9	2260	11	AK035624	AK035624 Mus muscu
45	158	13.8	1886	11	AK004500	AK004500 Mus muscu

ALIGNMENTS

RESULT 1
AK014524
LOCUS
DEFINITION
AK014524 1889 bp mRNA linear HTC 05-DEC-2002
Mus musculus 0 day neonate skin cDNA, RIKEN full-length enriched library, clone:4631416M05 product:serine (or cysteine) proteinase inhibitor, clade B (ovalbumin), member 7, full insert sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AK014524.1 GI:12852434
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
2
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

REFERENCE
AUTHORS

3. Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Koino, H., Akiyama, J., Nishi, K., Kitsuwa, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Tashiro, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, X., Izawa, M., Obara, S., Watanuki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system—384-format sequencing pipeline with 384 multipipillary sequencer Genome Res. 10 (11), 1757-1771 (2000).

TITLE

FEATURES

```

location/Qualifiers
1. 1889
   /organism="Mus musculus"
   /mol_type="mRNA"
   /strain="C57BL/6J"
   /db_xref="FANTOM_DB:4631416M05"
   /db_xref="MGI:1903231"
   /db_xref="taxon:10090"
   /clone="4631416M05"
   /tissue_type="skin"
   /clone_lib="Riken full-length enriched mouse cDNA library"
   /dev_stage="0 day neonate"
66. 1208

```

52

PUBMED
REFERENCE

and Miyashita-Takahara, T.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409 (6821), 685-690 (2001)
MEDLINE 21085660
 11217851
 PUBMED

5
REFERENCE
AUTHORS
TITLE
JOURNAL
6
REFERENCE
AUTHORS

REFERENCE
AUTHORS

5
The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase 1 & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 1889)
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
Furuno, M., Hasegaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
Hiraoaka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
Kasakawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishii, K.,
Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
Saito, H., Saito, P., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
Shibara, K., Shibata, Y., Shingawa, A., Shiraki, T., Sogabe, Y.,
Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
Tejima, Y., Toyota, T., Yamamura, T., Yasunishi, A., Yoshida, K.,
Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-922).

COMMENT

document

PMID: 19429303, 19521919

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAGAGAGGACCAAGCTCTTTTCTTTTCTT-3']. cDNA was prepared by using rNaseless thermo-activated reverse transcriptase


```

Db      604 CAAATCTATGATGAGACCAATCCAGGCGATGACCTCTAAGTATGTTGAGGTAATTT 663
QY      826 TTTCTCAGTTCAAGATAGAGAAGATATGAAATGAACAATATTTGAGAGCCCTAGGG 885
Db      664 TTTCTCAGTTCAAGATAGAGAAGATATGAAATGAACAATATTTGAGAGCCCTAAGG 723
QY      886 CTGAAGATATC-TTTGATGATGATCCAAAGGAGATCTCTGGGATGCTTCCGGGGGCTG 944
Db      724 CTGAAGATATCTTTTATGATGATCCAAAGGAGATCTCTGGGATGCTTCCGGGGGCTG 783
QY      945 TCT 947
Db      784 TCT 786

RESULT 3
AW381394/c
LOCUS      PMO-HT0302-271099-001-e08 HT0302 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW381394
VERSION    AW381394.1 GI:6886053
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 638)
            HGCP http://www.ludwig.org.br/ORESTES.
            TITLE The FAPESP/LICR Human Cancer Genome Project
            JOURNAL Unpublished
            COMMENT Contact: Simpson A.J.G.
                    Laboratory of Cancer Genetics
                    Ludwig Institute for Cancer Research
                    Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
                    Brazil
                    Tel: +55-11-2704922
                    Fax: +55-11-2707001
                    Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PMO&tl2=PMO-HT0302-271099-001-e08&tl3=1999-10-27&tl4=1)
            Seg primer: puc 18 forward
            High quality sequence start: 12
            High quality sequence stop: 636.
FEATURES   source
            location/Qualifiers
            i..638
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /dev_stage="Adult"
            /clone_lib="HT0302"
            /note="Organ: head neck; Vector: puc18; Site 1: Smar;
            Site 2: Smar; A mini-library was made by cloning products
            derived from ORESTES PCR (U.S. Letters Patent application
            No. 196,716 - Ludwig Institute for Cancer Research)
            profiles into the pUC 18 vector. Reverse transcription of
            tissue mRNA and cDNA amplification were performed under
            low stringency conditions."
BASE COUNT 175 a 149 c 132 g 182 t
ORIGIN

Query Match 52.9%; Score 604.4; DB 9; Length 638;
Best Local Similarity 99.8%; Pred. No. 3.7e-154;
Matches 605; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      32 TTTGCTTCAACTGTTTCAAGAGATGATGACATCAAGGAATGGAATGTTCTTTT 91
Db      608 TTTGCTTCAACTGTTTCAAGAGATGATGACATCAAGGAATGGAATGTTCTTTT 549
QY      92 CCTCTGAGCCCTCTTGGTCCCTGGCCCTGGTCCCTGGCCCTCAAGATGATCCC 151
Db      548 CCTCTGAGCCCTCTTGGTCCCTGGCCCTGGTCCCTGGCCCTCAAGATGATCCC 489

```

```

QY      152 TCTCTCAGTTGATTAAGTTGCTTCATGTTTAACTGCTCAGGATATGAAACTTCTTA 211
Db      488 TCTCTCAGTTGATTAAGTTGCTTCATGTTTAACTGCTCAGGATATGAAACTTCTTA 429
QY      212 ATAGTCAGTCAGGGCTCCAGTCTCAACTGCAAAAGAGTTTTTCTGATATAAATCATCCC 271
Db      428 ATAGTCAGTCAGGGCTCCAGTCTCAACTGCAAAAGAGTTTTTCTGATATAAATCATCCC 369
QY      272 ACAAGGATTAATGATCTCAGCAATGTAATGGGCTTTTTTGGCTGAAAAAGTGTATGGCTTTT 331
Db      368 ACAAGGATTAATGATCTCAGCAATGTAATGGGCTTTTTTGGCTGAAAAAGTGTATGGCTTTT 309
QY      332 ATAGAGCTACATTTGAGTGTGCGGCAAAATATATACGATGCCAACGAAAGTGGAGCGAGTTGACT 391
Db      308 ATAGAGCTACATTTGAGTGTGCGGCAAAATATATACGATGCCAACGAAAGTGGAGCGAGTTGACT 249
QY      392 TTACGATCATTTTGAAGACACTAGACGTAATATTAATTAAGTGGGTGAAATGAACAC 451
Db      248 TTACGATCATTTTGAAGACACTAGACGTAATATTAATTAAGTGGGTGAAATGAACAC 489
QY      452 ATGCGAAAATCAAGAACTGATTTGGTGAAGTGGCATAAGCTCATCTGCTGTATATGATGC 511
Db      188 ATGCGAAAATCAAGAACTGATTTGGTGAAGTGGCATAAGCTCATCTGCTGTATATGATGC 129
QY      512 TGGTGAATGCTGTGTACTTCAAGGCAAGTGGCAATCAGCTTCACCAAGAGCGAACCAC 571
Db      128 TGGTGAATGCTGTGTACTTCAAGGCAAGTGGCAATCAGCTTCACCAAGAGCGAACCAC 69
QY      572 TAAATTGGCAATTTCAAAATCTCCCAAGTCTCTGGGAAGGCGAGTCGCCATGATGCATCAGG 631
Db      68 TAAATTGGCAATTTCAAAATCTCCCAAGTCTCTGGGAAGGCGAGTCGCCATGATGCATCAGG 9
QY      632 AACGGA 637
Db      8 AACGGA 3

RESULT 4
AW381390/c
LOCUS      PMO-HT0302-271099-001-d11 HT0302 Homo sapiens linear EST 04-FEB-2000
DEFINITION PMO-HT0302-271099-001-d11 HT0302 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW381390
VERSION    AW381390.1 GI:6886049
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 638)
            HGCP http://www.ludwig.org.br/ORESTES.
            TITLE The FAPESP/LICR Human Cancer Genome Project
            JOURNAL Unpublished
            COMMENT Contact: Simpson A.J.G.
                    Laboratory of Cancer Genetics
                    Ludwig Institute for Cancer Research
                    Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
                    Brazil
                    Tel: +55-11-2704922
                    Fax: +55-11-2707001
                    Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PMO&tl2=PMO-HT0302-271099-001-d11&tl3=1999-10-27&tl4=1)
            Seg primer: puc 18 forward
            High quality sequence start: 30
            High quality sequence stop: 562.
FEATURES   source
            location/Qualifiers
            i..638
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"

```



```

Db      549  GGCAAGATTCTTCCCTTGA 569
|||||
B1871335      824 bp  mRNA  linear  EST 11-OCT-2001
B1871335      603392761F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5402566 5',
DEFINITION    mRNA sequence.
ACCESSION     B1871335
VERSION       B1871335.1  GI:16045010
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS       Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE         1 (bases 1 to 824)
JOURNAL       NIH-MGC http://mgc.nci.nih.gov/.
COMMENT       National Institutes of Health, Mammalian Gene Collection (MGC)
              Contact: Robert Strausberg, Ph.D.
              Email: rsb@nih.gov
              Tissue Procurement: ATCC
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Icyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLAM12027 row: 1 column: 15
              High quality sequence stop: 824.
              Location/Qualifiers
                1..824
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:5402566"
                  /tissue types="adenocarcinoma, cell line"
                  /lab_hosts="DH10B (phage-resistant)"
                  /clone_lib="NIH_MGC_90"
                  /note="Organ: liver; Vector: pCWV-SF0RT6; Site 1: NotI;
                  Site 2: SalI; Cloned unidirectionally; oligo-dr primed.
                  Average insert size 1.7 kb. Library enriched for
                  full-length clones and constructed by Life Technologies.
                  Note: this is a NIH_MGC Library."
              232 a 186 c 164 g 242 t

BASE COUNT  232 a 186 c 164 g 242 t
ORIGIN
Query Match  46.9%; Score 536; DB 12; Length 824;
Best Local Similarity 100.0%; Pred. No. 2.3e-135;
Matches 536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  608  AGGCACTGCCCATGATGATCAGGAACGGAAGTCAATTGCTCTTTTATGAGACCAT 667
Db      |||||
12  AGGCACTGCCCATGATGATCAGGAACGGAAGTCAATTGCTCTTTTATGAGACCAT 71
QY      |||||
668  CAATGAAGATCTTGGAGTCAGATCAATGATGGTGCATAAACAATGATGCTGCTGCTG 727
Db      |||||
72  CAAATGAGATCTTGGAGTCAGATCAATGATGGTGCATAAACAATGATGCTGCTGCTG 131
QY      |||||
728  AGAATGACCTCTGAAATGAAACAAATGACCTTTGAGATCTAATGGAATGGAACA 787
Db      |||||
132  AGAATGACCTCTGAAATGAAACAAATGACCTTTGAGATCTAATGGAATGGAACA 191
QY      |||||
788  ATCCAGGCGGATGACCTCTAAGTATGTTGAGGTATTTTCTCAGTCTCAATGAGA 847
Db      |||||
192  ATCCAGGCGGATGACCTCTAAGTATGTTGAGGTATTTTCTCAGTCTCAATGAGA 251
QY      |||||
848  AGAATTATGAATGAAACAATATTTGAGAGCCCTAGGCTGGAAGATATCTTTGATGAT 907
Db      |||||
252  AGAATTATGAATGAAACAATATTTGAGAGCCCTAGGCTGGAAGATATCTTTGATGAT 311
QY      |||||
908  CCAAGACAGATCTCTCTGGGATTGCTTCGGGGGGTCTCTGTATATATCAAGGATGATGC 967
|||||
312  CCAAGACAGATCTCTCTGGGATTGCTTCGGGGGGTCTCTGTATATATCAAGGATGATGC 371
|||||
968  ACAAATCTTACATAGAGGTCACTGAGAGGCGCAGGAGGCTACTGCTGCCACAGAACTA 1027
Db      |||||
372  ACAAATCTTACATAGAGGTCACTGAGAGGCGCAGGAGGCTACTGCTGCCACAGAACTA 431
QY      |||||
1028  ATATTGTAGAAAGCACTCTCCTCAGTCCAGCTGTTTGTAGAGTCAACACCATCTCTAT 1087
Db      |||||
432  ATATTGTAGAAAGCACTCTCCTCAGTCCAGCTGTTTGTAGAGTCAACACCATCTCTAT 491
QY      |||||
1088  TTGTTATCAGGAGCATGACATCATCTTATTCAGTGCACAAAGTTCTTGCCCTTGA 1143
Db      |||||
492  TTGTTATCAGGAGGATGACATCATCTTATTCAGTGCACAAAGTTCTTGCCCTTGA 547

RESULT 7
BY713938
LOCUS       BY713938
DEFINITION RIKEN full-length enriched, 0 day neonate skin Mus
ACCESSION   BY713938
VERSION     BY713938.1  GI:27126108
KEYWORDS    Mus musculus (house mouse)
SOURCE      Mus musculus
ORGANISM    Mus musculus
REFERENCE   1 (bases 1 to 945)
AUTHORS     Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
            Nikaide,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I., Kiyosawa,H.,
            Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C.,
            Gojibori,T., Baldarelli,R., Hill,D.P., Bult,C., Hume,D.A.,
            Quackenbush,J., Schriml,E.M., Kapapin,A., Matsuda,H., Batalov,S.,
            Beisel,X.W., Blake,J.A., Bradt,D., Brusic,V., Choche,C., Corbani,
            L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.F., Forrest,
            A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Gozdek,A.,
            Gough,J., Grimmond,S., Gustincich,S., Hirokawa,N., Jackson,I.J.,
            Jarvis,E.D., Kanaya,A., Kawaji,H., Kawasawa,Y., Kedzierski,R.M.,
            King,B.J., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,
            P.A., Maglott,D.R., Maltai,L., Marchionni,L., McKenzie,L., Mikl,
            H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Perleag,G.,
            Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D.,
            Panachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,
            B.Z., Ringwald,M., Sandelin,A., Schneider,C., Sempel,C.A., Setou,
            M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,
            Watanabe,Y., Wells,C., Wilming,I.G., Wyszaw-Boris,A., Yanagisawa,
            M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A.,
            Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,
            M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
            Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,
            Y., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,
            K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,
            E.S., Rogers,J., Birney,E. and Hayashizaki,Y.
            Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
            Nature 420, 563-573 (2002)
JOURNAL     Nature
MEDLINE     22354683
PUBMED      12466851
COMMENT     Contact: Yoshinide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center (GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@sc.riken.go.jp
            URL: http://genome.gsc.riken.go.jp/
            Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,
            S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F., Imotani,K.,
            Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y., Kondo,S., Konno,
            H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K.,

```


source

1...498
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:9980410603 ; IMAGE:4749599"
 /tissue_type="squamous cell carcinoma"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NCI_CGAP_Skn4"
 /note="Organ: Skin; Vector: pCMV-SPORT6; Site:1: NotI;
 Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.5kb. Library constructed by Life
 Technologies. Note: this is a NCI_CGAP Library."
 BASE COUNT 153 a 101 c 114 g 130 t
 ORIGIN

Query Match 43.4%; Score 456.4; DB 13; Length 498;
 Best Local Similarity 99.8%; Pred. No. 1.2e-124;
 Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 583 TTCAATCTCCCAAGTCTCTGGGAAGCAGTCCCATGATGATCAGGACGGAAGTTC 642
 Db 1 TTCAATCTCCCAAGTCTCTGGGAAGCAGTCCCATGATGATCAGGACGGAAGTTC 60
 QY 643 AATTCTCTGTATTGAGGACCATCAATGAAGATCTTGGCTCAGATCAATGTGGC 702
 Db 61 AATTCTCTGTATTGAGGACCATCAATGAAGATCTTGGCTCAGATCAATGTGGC 120
 QY 703 ATAAACATGAGTCTCTGGCTGAGATGACCTCTGAAATGAAACAAACTGACC 762
 Db 121 ATAAACATGAGTCTCTGGCTGAGATGACCTCTGAAATGAAACAAACTGACC 180
 QY 763 TTTCAGAACTAATGAGATGACCAATCCAGGCGATGACCTCTAAGTATGTAGSTA 822
 Db 181 TTTCAGAACTAATGAGATGACCAATCCAGGCGATGACCTCTAAGTATGTAGSTA 240
 QY 923 TTTTTCCTCAGTTCAGATGAGAGAAATATGAATGAACAAATATTTGAGAGCCTA 882
 Db 241 TTTTTCCTCAGTTCAGATGAGAGAAATATGAATGAACAAATATTTGAGAGCCTA 300
 QY 883 GGCTGAAGATATCTTTGATGATCCAAAGCAGATCTCTGGGATGCTCGGGGT 942
 Db 301 GGCTGAAGATATCTTTGATGATCCAAAGCAGATCTCTCGGGATGCTCGGGGT 360
 QY 943 CGTCTGATATATCAGGATGACAAATCTTACATAGGATCAGTGGAGGACCC 1002
 Db 361 CGTCTGATATATCAGGATGACAAATCTTACATAGGATCAGTGGAGGACCC 420
 QY 1003 GAGCTACTGCTGCCAGGAGTAATATTTGAGAAAGCAACTCCCTCAGTCCAGCTG 1062
 Db 421 GAGCTACTGCTGCCAGGAGTAATATTTGAGAAAGCAACTCCCTCAGTCCAGCTG 480
 QY 1063 TTTAGAGTGCACCCCA 1080
 Db 481 TTTAGAGTGCACCCCA 498

RESULT 9

CB590890
 LOCUS CB590890
 DEFINITION AGENCOURT_12567930 NIH_MGC_136 Mus musculus cDNA clone
 IMAGE:30288309 5', mRNA sequence.
 ACCESSION CB590890
 VERSION CB590890.1 GI:29508746
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 906)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. David Rowe
 cDNA Library Preparation: Invitrogen Corp
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: NDAM325 row: 1 column: 22
 High quality sequence stop: 668.

FEATURES
source

1...906
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:30288309"
 /tissue_type="embryonic limb, maxilla and mandible"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_136"
 /note="Vector: pCMV-SPORT6.1.cdb; Site:1: EcoRV; Site:2:
 NotI; Normalized, full-length enriched library from pool
 of mouse embryonic limb, maxilla and mandible (5, 4 and 1 limb and
 day 17.5, 18.5 and newborn (mandible (5, 4 and 1 limb and
 jaw equivalents from respective days). Cloned
 directionally, oligo-dT primed
 (5'-GACTAGTCTTGTAGTCGAGCGCCGCC (T)15-3'. Size selected
 for the >1kb fragments, average insert size 1.2 kb.
 Normalization to Cot 7.5. Tissue contributed by David
 Rowe; library constructed by ResGen, Invitrogen Corp.
 Note: this is a NIH MGC Library."
 BASE COUNT 285 a 186 c 204 g 231 t
 ORIGIN

Query Match 39.5%; Score 451.2; DB 14; Length 906;
 Best Local Similarity 79.5%; Pred. No. 3.9e-112;
 Matches 534; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

QY 1 ATGGCTCCCTTCTGCGAAGATGAGATGCTTCTCTCTGAGGCTCTGCTGCGCCGCGC 60
 Db 54 ATGGCTCCCTTCTGCGAAGATGAGATGCTTCTCTCTGAGGCTCTGCTGCGCCGCGC 113
 QY 61 GACATCAAGGAATGGAATGTGTTTCTCTCTGAGGCTCTGCTGCGCCGCGC 120
 Db 114 AGTAGCAAGGAATGGAATGTGTTTCTCTCTGAGGCTCTGCTGCGCCGCGC 173
 QY 121 CTGGTCCGTTGGGGCTCAAGATGATCCCTCTCTCAGATTGATAAGTTCCTTCATGTT 180
 Db 174 CTAATCCGTTGGGGCTCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 233
 QY 181 AACACTGCTCAGGATGAGAACTCTCTAATAGTCTGAGGCTCCAGTCTCACTG 240
 Db 234 AACATACCATCAAGCAAGGAATCTCTAATATCAAGCAAGCAAGCAAGCAAGCAAG 293
 QY 241 AAAAGAGTTTTTCTGATATAAATGATCCCAAGGATTATGATCTCAGATTGTGTAAT 300
 Db 294 AAAAGAGTTTTTCTGATATAAATGATCCCAAGGATTATGATCTCAGATTGTGTAAT 353
 QY 301 GGGCTTTTCTGTAAGGATGATGCTTTCTAAGGATCTAATGAGTGTGCGGAAAAA 360
 Db 354 GGAGTTTTTGGCAGAAAAAGTCTGATCTCTAAGAACTACATTTGAGTGTGCTGAAAC 413
 QY 361 TTATACGATGCCAAGTGGAGGCTGCTGCTTACGAATCTTTAGACACTAGACCT 420
 Db 414 TTATACGATGCCAAGTGGAGGCTGCTGCTTACGAATCTTTAGACACTAGACCT 473
 QY 421 AATATTAAATAGTGGGTTGAAATGAAACATGCGCAATCAAGCAAGCTGATGTTGTA 480
 Db 474 AATATTAAATAGTGGGTTGAAATGAAACATGCGCAATCAAGCAAGCTGATGTTGTA 533
 QY 481 GGTGGCATAGCTCATCTGCTGTATGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 540
 Db 534 AGCAGCTCAGCTCGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 593

QY 541 TGGCAATCAGCTTCCACGAGGACGACATTAATTCGCAATTCCTCAATCTCCAGTGC 600
 |||||
 Db 594 TGGAAATCGGCTTCCACAGACTGATACCTCAGTCCGCTTTAGGTCTCCACGCTGT 653
 |||||
 QY 601 TCTGGGAAGCAGTCCGATGATGATCAGGAGGAGGAGTTCGAATTTGCTCTTATTGAG 660
 |||||
 Db 654 CTTGGAAAGTAGTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 713
 |||||
 QY 661 GACCCATCAATG 672
 |||||
 Db 714 CAGCCACCAATG 725
 |||||

RESULT 10
 BG319597

LOCUS 405 bp mRNA linear EST 27-FEB-2001

DEFINITION 4dn19 Human Keratinocyte Subtraction Library- Downregulated Transcripts Homo sapiens cDNA similar to serine (cysteine) proteinase inhibitor, class B (ovalbumin), member 7 (SERPINB7), mRNA sequence.

ACCESSION BG319597 GI:13129272

VERSION BG319597.1

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 405)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Schlager, J.J., Benjamin, H.R., Ali, K., Levine, C.F., Dodds, A.D., Avery, D.P., Clark, J.H. and Hofmann, C.M.

TITLE Human Epidermal Keratinocyte Subtraction Library-Downregulated Transcripts

JOURNAL Unpublished

COMMENT Contact: Schlager, J.J.
 Molecular Toxicology
 United States Army Medical Research Institute of Chemical Defense (USAMRICD)
 3100 Ricketts Point Road; ATTN: WCMR-UV-PA; APG-EA, MD 21010-5400, USA
 Tel: 410 436 1940
 Fax: 410 436 1960
 Email: John.Schlager@AMEDD.ARMY.MIL.

FEATURES
 source
 1. 405
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /sex="female"
 /tissue_type="skin"
 /cell_type="primary keratinocyte"
 /cell_line="Clonetics 4075"
 /dev_stage="adult"
 /lab_host="E. coli (ampicillin-resistant)"
 /clone_lib="Human Keratinocyte Subtraction Library-Downregulated Transcripts"
 /note="Vector: pT-Adv; Library preparation: Sequence isolated using PCR-select (Clontech) subtraction library construction (Diatchenko et al., (1996) Proc. Natl. Acad. Sci. USA 93: 6025-6030) after sulfur mustard exposure of primary human adult epidermal keratinocytes. Subtraction library cloned for random sequence selection into pT-Adv vector."

BASE COUNT 124 a 81 c 90 g 110 t

ORIGIN

Query Match 35.4%; Score 405; DB 10; Length 405;
 Best Local Similarity 100.0%; Pred. No. 1.2e-99;
 Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 714 CTTCTGCTGCTGAGTATGACCTCTCTGAAATGAAACAAACTGACCTTTGAGATCT 773
 |||||
 Db 1 CTTCTGCTGCTGAGATGACCTCTCTGAAATGAAACAAACTGACCTTTGAGATCT 60
 |||||

QY 774 AATGGAATGGACCAATCCAGGCGAATGACCTCTAAGTATGTTGAGTATTTTTCCTCA 833
 |||||
 Db 61 AATGGAATGGACCAATCCAGGCGAATGACCTCTAAGTATGTTGAGTATTTTTCCTCA 120
 |||||
 QY 834 GTTCAAGATAGAGAAGATATGAAATGAAACAAATATTGAGACCCCTAGGGCTGAAAGA 893
 |||||
 Db 121 GTTCAAGATAGAGAAGATATGAAATGAAACAAATATTGAGACCCCTAGGGCTGAAAGA 180
 |||||
 QY 894 TATCTTTGATGAATCCAAAGCAGATCTCTGGGATTCCTTCGGGGGGTCTCTGTATAT 953
 |||||
 Db 181 TATCTTTGATGAATCCAAAGCAGATCTCTCTGGGATTCCTTCGGGGGGTCTCTGTATAT 240
 |||||
 QY 954 ATCAAGGATGATGACAAATCTTACATAGAGTCACTGAGGAGGCGACCCGAGCTACTGCG 1013
 |||||
 Db 241 ATCAAGGATGATGACAAATCTTACATAGAGTCACTGAGGAGGCGACCCGAGCTACTGCG 300
 |||||
 QY 1014 TCCACAGGAGTATATTGTTAGAAAAGCAATCCCTGATGTCACCGCTGTTTAGAGCTGA 1073
 |||||
 Db 301 TCCACAGGAGTATATTGTTAGAAAAGCAATCCCTGATGTCACCGCTGTTTAGAGCTGA 360
 |||||

QY 1074 CCACCCATCTCTATTGTTTATCAGGAGGATGACATCATCTTATT 1118
 |||||
 Db 361 CCACCCATCTCTATTGTTTATCAGGAGGATGACATCATCTTATT 405
 |||||

RESULT 11
 BF247923

LOCUS 627 bp mRNA linear EST 14-NOV-2000

DEFINITION 601859005F1 NIH_MGC_58 Homo sapiens cDNA clone IMAGE:4069704 5', mRNA sequence.

ACCESSION BF247923

VERSION BF247923.1

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 627)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.
 Email: c9apbs@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: CLONETECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LHCW912 row: i column: 01
 High quality sequence stop: 385.
 Location/Qualifiers
 1. 627
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4069704"
 /tissue_type="hypertrophoma"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NIH_MGC_58"
 /note="Organ: kidney; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggccctcgcc); Site_2: SfiI (ggccatgatgccc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CAGCGCATATGCCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGCGGCGGCATG-JT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.35 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

BASE COUNT 210 a 129 c 149 g 139 t

Qy 649 TCTGTATTAGGACCCATCAATGAAGATTCCTGAGCT-----CAGATACATGGT 699
Db |||||
727 GCTATCATAAAGGAGCCAGAGTGCAGTCTCTTGACCTTCCTATGCGCACACAACTA 786
Qy |||||
700 GGCATARACATGTACGTTCTGCTGAGATGACCTCTCTGAAATTTGAAAACAACTG 759
Db |||||
787 AGAATGATCATCTGCTGCTCTAGGCACACCCAGCGGTGAGCCAGATAGAAAACACCTG 846
Qy 760 ACCTTTCAGAACTTAATGGAATGGACCAATCCRAGCGCAATGACCTCTAAGTATGTTGAG 819
Db |||||
847 AATGTGAAGATGTTGAGAGTGGACCTTAACCCCTCTTAACATGGTGAAGAGAGTGGAT 906
Qy 820 GTAATTTTCTCAGTTCAAGATAGAGAGATTAATGAATGAACAATAATTTGAGAGCC 879
Db |||||
907 GTTCACATCCCTAAATTCAGTCTTTCAGTAAATAGCACCTAAACACCCCTATTGAAATCC 966
Qy 880 CTAGGGCTGAAGATATCTTTGATGATCCAAAGCAGATCTCTCTGGATTCCTTCGGGG 939
Db |||||
967 CTAGGCATGAGGACATCTTCAAGTACCAATGCTGATCTCTGGAATGTCACCGGAC 1026
Qy 940 GGTCTCTGTATATATCAAGGATGATGCACAAATCTTACATAGAGTCTCTCAGAGGGC 999
Db |||||
1027 AAGGCTATATTTATCTAAGTTGTTTCACAGTATATGTGGATGTCAATGAAGAGGC 1086
Qy 1000 ACCGAGCTACTCTGCCACAGGAAGTAATATTGTAGAAAAGCAACTCCCTCAGTCCACG 1059
Db |||||
1087 ACAGAGGACGCTCAGCCACTGGGGAGAGTATTCTGTGAGCGACTCCCTGTCACAGTT 1146
Qy 1060 CTGTTTAGAGCTGACCCCATTCCTATTGTTATCAGGAAGGA---TCACATCATCTTA 1116
Db |||||
1147 CAGTTCACAGCCATTTGCTTCTTCTGTTCTTTATCTGGATGAGTCTGGCAATATTTTA 1206
Qy 1117 TTCAGTGGCAAGTTTCTTGCCCTT 1141
Db |||||
1207 TTTCGTGGCAAGTTTGCTTCCAT 1231

Search completed: December 14, 2003, 03:19:22
Job time : 2782 secs

result	No.	Score	Query		ID	Description
			Match	Length		
1	380	100.0	380	2	US-08-472-659-34	Sequence 34, Appl
2	380	100.0	380	2	US-08-474-661-34	Sequence 34, Appl
3	380	100.0	380	2	US-08-611-977-34	Sequence 34, Appl
4	19	5.0	19	2	US-08-611-977-3	Sequence 3, Appl
5	18	4.7	20	2	US-08-611-977-6	Sequence 6, Appl
6	17	4.5	17	2	US-08-611-977-4	Sequence 4, Appl
7	17	4.5	20	2	US-08-472-659-6	Sequence 6, Appl
8	17	4.5	20	2	US-08-474-661-6	Sequence 6, Appl
9	13	3.4	382	1	US-07-768-286B-6	Sequence 6, Appl
10	13	3.4	382	1	US-08-487-823B-3	Sequence 6, Appl
11	13	3.4	382	2	US-08-997-040-3	Sequence 3, Appl
12	13	3.4	382	2	US-09-203-237-3	Sequence 3, Appl
13	13	3.4	392	1	US-07-768-286B-4	Sequence 3, Appl
14	13	3.4	405	1	US-08-421-714-7	Sequence 4, Appl
15	13	3.4	405	1	US-08-477-112-7	Sequence 7, Appl
16	13	3.4	405	2	PCT-US93-08332-7	Sequence 7, Appl
17	13	3.4	415	1	US-07-911-531-19	Sequence 19, Appl
18	13	3.4	415	1	US-07-593-636A-19	Sequence 19, Appl
19	13	3.4	415	1	US-07-768-286B-2	Sequence 2, Appl
20	13	3.4	415	1	US-09-026-408-4	Sequence 4, Appl
21	13	3.4	434	1	US-07-679-052A-15	Sequence 15, Appl
22	13	3.4	438	1	US-07-679-052A-17	Sequence 17, Appl
23	13	3.4	438	1	US-08-472-659-4	Sequence 4, Appl
24	11	2.9	17	2	US-08-474-661-4	Sequence 4, Appl
25	11	2.9	17	2	US-08-611-977-9	Sequence 9, Appl
26	10	2.6	10	2	US-08-472-659-3	Sequence 3, Appl
27	10	2.6	19	2	US-08-472-659-3	Sequence 3, Appl

;
; LENGTH: 380 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-472-659-34

Query Match 100.0%; Score 380; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASLAANAEEFCNLFREMDNQNGNVFFSSLSLFAALALVRLGAQDDSLSQIDKLLHV 60
Db 1 MASLAANAEEFCNLFREMDNQNGNVFFSSLSLFAALALVRLGAQDDSLSQIDKLLHV 60
QY 61 NTASGYNSNSQSGQLKRVFSDINASHKDYDLSIVNGLFAEKVYGFHKVIECAEK 120
Db 61 NTASGYNSNSQSGQLKRVFSDINASHKDYDLSIVNGLFAEKVYGFHKVIECAEK 120
QY 121 LYDAKVERVDFTNHLEDFRRNKVVENETHGKIKNVIGGGISSAVVWLVNAVYFKGK 180
Db 121 LYDAKVERVDFTNHLEDFRRNKVVENETHGKIKNVIGGGISSAVVWLVNAVYFKGK 180
QY 181 WQSAFTKSETINCHFKSPKSGKAVAMHQERKNLSVIEDPSMKILRLYNGGINMYVL 240
Db 181 WQSAFTKSETINCHFKSPKSGKAVAMHQERKNLSVIEDPSMKILRLYNGGINMYVL 240
QY 241 LPENDLSEIENKLTQNLMEWTPREMTSKYVEVFPQPKIEKNVEMKQYLRALGLKIDIF 300
Db 241 LPENDLSEIENKLTQNLMEWTPREMTSKYVEVFPQPKIEKNVEMKQYLRALGLKIDIF 300
QY 301 DESKADLSGIASGRLYISRMHKSIVTEEGTEATAATGNSNIVEKQLPOSTLFRADHP 360
Db 301 DESKADLSGIASGRLYISRMHKSIVTEEGTEATAATGNSNIVEKQLPOSTLFRADHP 360
QY 361 FLFVIRKDDIILFSGKVSCP 380
Db 361 FLFVIRKDDIILFSGKVSCP 380

RESULT 2

US-08-474-661-34
; Sequence 34, Application US/08474661
; Patent No. 5874253
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Ruyuki
; APPLICANT: TSURUOKA, No. 5874253uo
; APPLICANT: NAKAZATO, Hiroshi
; APPLICANT: MIURA, Kenju
; APPLICANT: ISHIDA, No. 5874253uhiro
; APPLICANT: KURIHARA, Tatsuya
; APPLICANT: YAMAICHI, Kozo
; APPLICANT: YAMAGUCHI, No. 5874253omi
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: George Mason Bldg., Washington & Prince Sts.
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,661
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/091,028

;
; FILING DATE: 14-JUL-1993
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5-067339
; FILING DATE: 04-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: REA, TERESA STANEX
; REGISTRATION NUMBER: 30,427
; REFERENCE/DOCKET NUMBER: 001560-204
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-6620
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-474-661-34

Query Match 100.0%; Score 380; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASLAANAEEFCNLFREMDNQNGNVFFSSLSLFAALALVRLGAQDDSLSQIDKLLHV 60
Db 1 MASLAANAEEFCNLFREMDNQNGNVFFSSLSLFAALALVRLGAQDDSLSQIDKLLHV 60
QY 61 NTASGYNSNSQSGQLKRVFSDINASHKDYDLSIVNGLFAEKVYGFHKVIECAEK 120
Db 61 NTASGYNSNSQSGQLKRVFSDINASHKDYDLSIVNGLFAEKVYGFHKVIECAEK 120
QY 121 LYDAKVERVDFTNHLEDFRRNKVVENETHGKIKNVIGGGISSAVVWLVNAVYFKGK 180
Db 121 LYDAKVERVDFTNHLEDFRRNKVVENETHGKIKNVIGGGISSAVVWLVNAVYFKGK 180
QY 181 WQSAFTKSETINCHFKSPKSGKAVAMHQERKNLSVIEDPSMKILRLYNGGINMYVL 240
Db 181 WQSAFTKSETINCHFKSPKSGKAVAMHQERKNLSVIEDPSMKILRLYNGGINMYVL 240
QY 241 LPENDLSEIENKLTQNLMEWTPREMTSKYVEVFPQPKIEKNVEMKQYLRALGLKIDIF 300
Db 241 LPENDLSEIENKLTQNLMEWTPREMTSKYVEVFPQPKIEKNVEMKQYLRALGLKIDIF 300
QY 301 DESKADLSGIASGRLYISRMHKSIVTEEGTEATAATGNSNIVEKQLPOSTLFRADHP 360
Db 301 DESKADLSGIASGRLYISRMHKSIVTEEGTEATAATGNSNIVEKQLPOSTLFRADHP 360
QY 361 FLFVIRKDDIILFSGKVSCP 380
Db 361 FLFVIRKDDIILFSGKVSCP 380

RESULT 3

US-08-611-977-34
; Sequence 34, Application US/08611977
; Patent No. 5972886
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Ruyuki
; APPLICANT: TSURUOKA, No. 5972886uo
; APPLICANT: NAKAZATO, Hiroshi
; APPLICANT: MIURA, Kenju
; APPLICANT: ISHIDA, No. 5972886uhiro
; APPLICANT: KURIHARA, Tatsuya
; APPLICANT: YAMAICHI, Kozo
; APPLICANT: YAMAGUCHI, No. 5972886omi
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404

CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/611,977
FILING DATE: 06-MAR-1996
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-611-977-34

Query Match 100.0%; Score 380; DB 2; Length 380;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASIAAANAEFCNLFREMDNNGNNGNFFSLSLFAALALVLAQDLSQIDKLHV 60
DB 1 MASIAAANAEFCNLFREMDNNGNNGNFFSLSLFAALALVLAQDLSQIDKLHV 60
QY 61 NTASGYCNSNSQSGLSQKRVFSDINASHKDYDLSIVNGLFAKVGREKDYICAEK 120
DB 61 NTASGYCNSNSQSGLSQKRVFSDINASHKDYDLSIVNGLFAKVGREKDYICAEK 120
QY 121 LYDAKVERVDFTHLEDTNRNINKWVENETHGKIKNVIGGGTSSSAWVWVNAVYFKGK 180
DB 121 LYDAKVERVDFTHLEDTNRNINKWVENETHGKIKNVIGGGTSSSAWVWVNAVYFKGK 180
QY 181 WQSAFTKSTINCHFKPKSCGKAVAMHQRKXNLSVIEDPSMKILBRVNGGINMYVL 240
DB 181 WQSAFTKSTINCHFKPKSCGKAVAMHQRKXNLSVIEDPSMKILBRVNGGINMYVL 240
QY 241 LPENDLSEIENKLFQMLMEWTNPRMTSKVEVFPQFKIKKNYEMQYLRALGKIDP 300
DB 241 LPENDLSEIENKLFQMLMEWTNPRMTSKVEVFPQFKIKKNYEMQYLRALGKIDP 300
QY 301 DESKADLSIASGGRLYISRMHKSIVETEGTEATAATGNSNIVEKQLPOSTLFRADHP 360
DB 301 DESKADLSIASGGRLYISRMHKSIVETEGTEATAATGNSNIVEKQLPOSTLFRADHP 360
QY 361 FLVIRKDDIILFSGKVSCP 380
DB 361 FLVIRKDDIILFSGKVSCP 380

RESULT 4

US-08-611-977-3

; Sequence 3, Application US/08611977

Patent No. 5972886
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURUOKA, No. 5972886uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5972886uhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAUCHI, Kozo
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/611,977
FILING DATE: 06-MAR-1996
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-611-977-3

Query Match 5.0%; Score 19; DB 2; Length 19;

Best Local Similarity 100.0%; Pred. No. 2.6e-11;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 126 VERVDFTHLEDTNRNINK 144

DB 1 VERVDFTHLEDTNRNINK 19

RESULT 5

US-08-611-977-6

; Sequence 6, Application US/08611977

; Patent No. 5972886

; GENERAL INFORMATION:

; APPLICANT: TSUJIMOTO, Masafumi

; APPLICANT: IWASA, Fuyuki

; APPLICANT: TSURUOKA, No. 5972886uo

; APPLICANT: NAKAZATO, Hiroshi

; APPLICANT: MIURA, Kenju

```

; APPLICANT: ISHIDA, No. 5972886uhiro
; APPLICANT: KURIHARA, Tatsuya
; APPLICANT: YAMAUCHI, Kozo
; APPLICANT: YAMAGUCHI, No. 5972886cmi
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611,977
; FILING DATE: 06-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; APPLICATION DATA:
; APPLICATION NUMBER: JP 6-067339
; FILING DATE: 04-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-204
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Peptide
; US-08-611-977-6

Query Match 4.7%; Score 18; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.5e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 305 ADLSGIASGGLYISRM 322
Db 1 ADLSGIASGGLYISRM 18

RESULT 6
US-08-611-977-4
; Sequence 4, Application US/08611977
; Patent No. 5972886
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Fuyuki
; APPLICANT: TSURUOKA, No. 5972886uo
; APPLICANT: NAKAZATO, Hiroshi
; APPLICANT: MIURA, Kenju
; APPLICANT: ISHIDA, No. 5972886uhiro
; APPLICANT: KURIHARA, Tatsuya
; APPLICANT: YAMAUCHI, Kozo
; APPLICANT: YAMAGUCHI, No. 5972886cmi
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611,977
; FILING DATE: 06-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; APPLICATION DATA:
; APPLICATION NUMBER: JP 6-067339
; FILING DATE: 04-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-204
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-611-977-4

Query Match 4.5%; Score 17; DB 2; Length 17;
Best Local Similarity 100.0%; Pred. No. 2e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 325 SYIVTEGTEATAATG 341
Db 1 SYIVTEGTEATAATG 17

RESULT 7
US-08-472-659-6
; Sequence 6, Application US/08472659
; Patent No. 5831030
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Fuyuki
; APPLICANT: TSURUOKA, No. 5831030uo
; APPLICANT: NAKAZATO, Hiroshi
; APPLICANT: MIURA, Kenju
; APPLICANT: ISHIDA, No. 5831030uhiro
; APPLICANT: KURIHARA, Tatsuya
; APPLICANT: YAMAUCHI, Kozo
; APPLICANT: YAMAGUCHI, No. 5831030cmi
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:

```

```

/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/472.659
/ FILING DATE: 07-JUN-1995
/ CLASSIFICATION: 536
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: JP 4-212305
/ FILING DATE: 17-JUL-1992
/ APPLICATION NUMBER: JP 5-067339
/ FILING DATE: 04-MAR-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/091,028
/ FILING DATE: 14-JUL-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: McGowan, Malcolm K.
/ REGISTRATION NUMBER: 39,300
/ REFERENCE/DOCKET NUMBER: 001560-248
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (703) 836-6620
/ TELEFAX: (703) 836-6620
/ INFORMATION FOR SEQ ID NO: 6:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 20 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ INFORMATION FOR SEQ ID NO: 6:
US-08-472-659-6

Query Match 4.5%; Score 17; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.3e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 305 ADLSGIASGGRLYISRM 321
DB 1 ADLSGIASGGRLYISRM 17

RESULT 8
US-08-474-661-6
/ Sequence 6, Application US/08474661
/ Patent No. 5874253
/ GENERAL INFORMATION:
/ APPLICANT: TSUJIMOTO, Masafumi
/ APPLICANT: IWASA, Fuyuki
/ APPLICANT: TSURUOKA, No. 5874253uo
/ APPLICANT: NAKAZATO, Hiroshi
/ APPLICANT: MIURA, Kenju
/ APPLICANT: ISHIDA, No. 5874253uhiro
/ APPLICANT: KURIHARA, Tatsuya
/ APPLICANT: YAWAICHI, Kozo
/ APPLICANT: YAMAGUCHI, No. 5874253omi
/ TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
/ NUMBER OF SEQUENCES: 34
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Burns, Doane, Swecker & Mathis
/ STREET: George Mason Bldg., Washington & Prince Sts.
/ CITY: Alexandria
/ STATE: Virginia
/ COUNTRY: United States
/ ZIP: 22313-1404
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/474.661
/ FILING DATE: 07-JUN-1995

```

```

/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/091,028
/ FILING DATE: 14-JUL-1993
/ APPLICATION NUMBER: JP 4-212305
/ FILING DATE: 17-JUL-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: JP 5-067339
/ FILING DATE: 04-MAR-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: REA, TERESA STANEK
/ REGISTRATION NUMBER: 30,427
/ REFERENCE/DOCKET NUMBER: 001560-204
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (703) 836-6620
/ TELEFAX: (703) 836-6620
/ INFORMATION FOR SEQ ID NO: 6:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 20 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ INFORMATION FOR SEQ ID NO: 6:
US-08-474-661-6

Query Match 4.5%; Score 17; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.3e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 305 ADLSGIASGGRLYISRM 321
DB 1 ADLSGIASGGRLYISRM 17

RESULT 9
US-07-768-286B-6
/ Sequence 6, Application US/07768286B
/ Patent No. 544153
/ GENERAL INFORMATION:
/ APPLICANT: GOSS, Neil H.
/ APPLICANT: RICHARDSON, Michael A.
/ TITLE OF INVENTION: VARIANTS OF PAI-2
/ NUMBER OF SEQUENCES: 21
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Foley & Lardner
/ STREET: 1800 Diagonal Road, Suite 500
/ CITY: Alexandria
/ STATE: Virginia
/ COUNTRY: USA
/ ZIP: 22313-0299
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07/768,286B
/ FILING DATE: 19911011
/ CLASSIFICATION: 514
/ PRICE APPLICATION DATA:
/ APPLICATION NUMBER: PCT/AU90/00603
/ FILING DATE: 20-DEC-1990
/ ATTORNEY/AGENT INFORMATION:
/ NAME: BENT, Stephen A.
/ REGISTRATION NUMBER: 23,768
/ REFERENCE/DOCKET NUMBER: 16786/157 CHAC
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (703) 836-9300
/ TELEFAX: (703) 683-4109
/ TELEX: 899149
/ INFORMATION FOR SEQ ID NO: 6:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 382 amino acids

```

; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-768-286B-6

Query Match 3.4%; Score 13; DB 1; Length 382;
Best Local Similarity 100.0%; Pred. No. 0.00024;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 MVLNNAVYFKGW 181
| | | | | | | | | |
Db 163 MVLNNAVYFKGW 175

RESULT 10

US-08-487-823B-3
; Sequence 3, Application US/08487823B
; Patent No. 5700924
; GENERAL INFORMATION:
; APPLICANT: Braxton, Scott M.
; APPLICANT: Diep, Dinh
; APPLICANT: Stuart, Susan G.
; TITLE OF INVENTION: NOVEL SERPIN DERIVED FROM HUMAN
; TITLE OF INVENTION: HYPOTHALAMUS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,823B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Luther, Barbara J.
; REGISTRATION NUMBER: 33,954
; REFERENCE/DOCKET NUMBER: PF-0039 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-852-0195
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 382 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
US-08-487-823B-3

Query Match 3.4%; Score 13; DB 1; Length 382;
Best Local Similarity 100.0%; Pred. No. 0.00024;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 MVLNNAVYFKGW 181
| | | | | | | | | |
Db 163 MVLNNAVYFKGW 175

RESULT 11

US-08-997-040-3
; Sequence 3, Application US/08997040
; Patent No. 5929210
; GENERAL INFORMATION:
; APPLICANT: Braxton, Scott M.
; APPLICANT: Diep, Dinh
; APPLICANT: Stuart, Susan G.
; TITLE OF INVENTION: NOVEL SERPIN DERIVED FROM HUMAN
; TITLE OF INVENTION: HYPOTHALAMUS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/997,040
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/487,823
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Luther, Barbara J.
; REGISTRATION NUMBER: 33,954
; REFERENCE/DOCKET NUMBER: PF-0039 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-852-0195
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 382 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
US-08-997-040-3

Query Match 3.4%; Score 13; DB 2; Length 382;
Best Local Similarity 100.0%; Pred. No. 0.00024;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 MVLNNAVYFKGW 181
| | | | | | | | | |
Db 163 MVLNNAVYFKGW 175

RESULT 12

US-09-203-237-3
; Sequence 3, Application US/09203237
; Patent No. 5955284
; GENERAL INFORMATION:
; APPLICANT: Braxton, Scott M.
; APPLICANT: Diep, Dinh
; APPLICANT: Stuart, Susan G.
; TITLE OF INVENTION: NOVEL SERPIN DERIVED FROM HUMAN
; TITLE OF INVENTION: HYPOTHALAMUS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
US-09-203-237-3

; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: US
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: PastSeq Version 1.5
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/203,237
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/997,040
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Luther, Barbara J.
 ; REGISTRATION NUMBER: 33,954
 ; REFERENCE/DOCKET NUMBER: PF-0039 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-855-0555
 ; TELEFAX: 415-852-0195
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 382 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; FRAGMENT TYPE: N-terminal
 ; ORIGINAL SOURCE:
 ; US-09-203-237-3

Query Match 3.4%; Score 13; DB 2; Length 382;
 Best Local Similarity 100.0%; Pred. No. 0.00024;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 MVLNNAVYFKGW 181
 Db 163 MVLNNAVYFKGW 175

RESULT 13
 US-07-768-286B-4
 ; Sequence 4, Application US/07768286B
 ; Patent No. 5444153
 ; GENERAL INFORMATION:
 ; APPLICANT: Goss, Neil H.
 ; APPLICANT: RICHARDSON, Michael A.
 ; TITLE OF INVENTION: VARIANTS OF PAI-2
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Foley & Lardner
 ; STREET: 1800 Diagonal Road, Suite 500
 ; CITY: Alexandria
 ; STATE: Virginia
 ; COUNTRY: USA
 ; ZIP: 22313-0299
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/768,2863
 ; FILING DATE: 1991101
 ; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/AU90/00603
 ; FILING DATE: 20-DEC-1990
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: BENT, Stephen A.
 ; REGISTRATION NUMBER: 29,768
 ; REFERENCE/DOCKET NUMBER: 16786/157 CHAC
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703)836-9300
 ; TELEFAX: (703)683-4109
 ; TELEX: 899149
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 392 amino acids
 ; TYPE: AMINO ACID
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-07-768-286B-4

Query Match 3.4%; Score 13; DB 1; Length 392;
 Best Local Similarity 100.0%; Pred. No. 0.00025;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 MVLNNAVYFKGW 181
 Db 173 MVLNNAVYFKGW 185

RESULT 14
 US-08-121-714-7
 ; Sequence 7, Application US/08121714
 ; Patent No. 5470970
 ; GENERAL INFORMATION:
 ; APPLICANT: Sager, Ruth
 ; TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH
 ; TUMOR SUPPRESSING ACTIVITY
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: U.S.A.
 ; ZIP: 02110-2804
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 ; COMPUTER: IBM PS/2 Model 502 or 55SX
 ; OPERATING SYSTEM: MS-DOS (Version 5.0)
 ; SOFTWARE: WordPerfect (Version 5.1)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/121,714
 ; FILING DATE:
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/938,823
 ; FILING DATE: 09/01/92
 ; APPLICATION NUMBER: 07/944,296
 ; FILING DATE: 02/28/92
 ; APPLICATION NUMBER: 07/662,216
 ; FILING DATE: 02/28/91
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fraser, Janis K.
 ; REGISTRATION NUMBER: 34,819
 ; REFERENCE/DOCKET NUMBER: 00530/072001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 542-5070
 ; TELEFAX: (617) 542-8906
 ; TELEX: 200154
 ; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 405
 ; TYPE: amino acid
 ; STRANDEDNESS:

Search completed: December 12, 2003, 16:27:58
Job time : 22 secs

```
; TOPOLOGY: linear
US-08-121-714-7

Query Match      3.4%; Score 13; DB 1; Length 405;
Best Local Similarity 100.0%; Pred.No. 0.00025;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      169 MVLNNAVYFKGW 181
Db      196 MVLNNAVYFKGW 208

RESULT 15
US-08-477-108A-7
; Sequence 7, Application US/08477108A
; Patent No. 5801001
; GENERAL INFORMATION:
; APPLICANT: Sager, Ruth
; APPLICANT: Zou, Zhigang
; APPLICANT: Anisowicz, Anthony
; TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH
; TITLE OF INVENTION: TUNOR SUPPRESSING ACTIVITY
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IEM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,108A
; FILING DATE: June 7, 1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/121,714
; FILING DATE: 09/01/93
; APPLICATION NUMBER: 07/938,823
; FILING DATE: 09/01/92
; APPLICATION NUMBER: 07/844,296
; FILING DATE: 02/28/92
; APPLICATION NUMBER: 07/662,216
; FILING DATE: 02/28/91
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 06570/002002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 405
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-477-108A-7

Query Match      3.4%; Score 13; DB 1; Length 405;
Best Local Similarity 100.0%; Pred.No. 0.00025;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      169 MVLNNAVYFKGW 181
Db      196 MVLNNAVYFKGW 208
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 16:27:28 : Search time 43 Seconds
(without alignments)
1402.700 Million cell updates/sec

Title: US-09-508-997A-2

Perfect score: 380

Sequence: 1 MASLAANAEEFCNLFREMD.....FLVIRKDDIILFSGKVSCT 380

Scoring table: OLIGO

Gapop 60.0, Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size: 0

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: A_Geneseq_19Jun03.*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	D3	ID	Description
1	380	100.0	380	15	AA198379	Human megakaryocyte
2	380	100.0	380	20	AA198254	Human megasin prote
3	380	100.0	380	21	AA198412	Human megasin prote
4	380	100.0	380	22	AA1983075	Human megasin prote
5	380	100.0	380	24	AA1985533	Lung cancer-associ
6	18	4.7	368	20	AA198256	Mouse megasin prote
7	18	4.7	368	21	AA1984151	Mouse megasin prote
8	18	4.7	368	22	AA1983077	Murine megasin prot
9	18	4.7	380	20	AA198255	Rat megasin protein

10	18	4.7	380	21	AA1984150	Rat megasin protein
11	18	4.7	380	22	AA1984286	Rat megasin protein
12	18	4.7	380	22	AA1983076	Rat megasin protein
13	17	4.5	20	15	AA1987112	Human megakaryocyte
14	16	4.2	16	20	AA198260	Megasin peptide 4.
15	16	4.2	16	21	AA1984145	Human megasin domain
16	16	4.2	16	21	AA1984146	Human megasin domain
17	15	3.9	15	20	AA198257	Megasin peptide 1.
18	15	3.9	15	20	AA198258	Megasin peptide 2.
19	15	3.9	15	20	AA198259	Megasin peptide 3.
20	15	3.9	15	21	AA1984143	Human megasin domain
21	15	3.9	15	21	AA1984144	Human megasin domain
22	15	3.9	15	22	AA1984147	Human megasin domain
23	15	3.9	15	22	AA1983078	Human megasin pepti
24	13	3.4	361	9	AA1980058	Sequence of plasm
25	13	3.4	382	12	AA1981214	Plasminogen activa
26	13	3.4	392	12	AA1981214	Plasminogen activa
27	13	3.4	415	8	AA1987030	Sequence of human
28	13	3.4	415	9	AA1980473	Sequence of human
29	13	3.4	415	10	AA1984160	Sequence of plasm
30	13	3.4	415	12	AA1980921	Encodes Plasminoge
31	13	3.4	415	12	AA19813007	Placental PAI-2
32	13	3.4	415	13	AA19820797	PAI-2 from pDBP3.
33	13	3.4	415	23	AA19872656	Human plasminogen
34	13	3.4	415	23	AA1984272	Plasminogen activa
35	13	3.4	415	24	AA19898694	Human PAI-2 polymo
36	13	3.4	423	22	AA19873764	Human colon cancer
37	13	3.4	434	12	AA19811352	PreA-plasminogen a
38	13	3.4	438	12	AA19811353	PreB-PAI-2 with al
39	11	2.9	17	15	AA19857110	Human megakaryocyte
40	10	2.6	16	23	AA19825330	Human LSI-01 prote
41	10	2.6	19	15	AA19857109	Human megakaryocyte
42	10	2.6	378	19	AA19830533	Human megakaryocyte
43	10	2.6	378	19	AA19830534	Human megakaryocyte
44	10	2.6	378	19	AA19830535	Human megakaryocyte
45	10	2.6	378	19	AA19830536	Human megakaryocyte

ALIGNMENTS

RESULT 1
AA198379
ID AA198379 standard; Protein; 380 AA.
XX
AC AA198379;
XX
DT 25-MAR-2003 (updated)
DT 15-AUG-1994 (first entry)
XX
DE Human megakaryocyte differentiation factor.
XX
KW Human megakaryocyte differentiation factor; MDF; thrombopoietin;
KW haematopoietic stimulating factor; thrombocytopaenia; platelet;
KW bone marrow transplantation; cancer chemotherapy.
OS Homo sapiens.
XX
PN EP1983884-A1.
XX
PD 23-FEB-1994.
XX
PF 19-JUL-1993; 93EP-0305654.
XX
PR 17-JUL-1992; 93JP-0212305.
PR 04-MAR-1993; 93JP-0067339.
XX
PA (SUNB) SUNTORY LTD.
PA (TSUJ) TSUJIMOTO M.
XX
PI Ishida N, Iwasa F, Kurihara T, Miura K, Nakazato H;
PI Tsujimoto M, Tsuruoka N, Yamaguchi N, Yamauchi K;
XX

```

DR WPI; 1994-050782/08.
DR N-PSDB; AAQ56670.
PT New megakaryocyte differentiation factor - isolated from human
PT epidermoid carcinoma cells, used to treat conditions involving a
PT decrease in platelets
XX
PS Claim 7; Page 30-32; 47pp; English.
XX
CC Human MDF can be isolated from a culture of human epidermoid
CC carcinoma A431 cells in protein-free medium. The MDF stimulates
CC differentiation of megakaryocytes from myeloid cells in the presence
CC of IL-3. The MDF acts in vivo as a thrombopoietin making it useful
CC for treatment of diseases involving a decrease in platelet number
CC (esp. thrombocytopaenia) such as occurs in bone marrow
CC transplantation and in chemotherapy. MDF has mol.wt. 55-57kd by
CC SDS-PAGE and exhibits an isoelectric point of 6.0-7.0.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 380 AA;
Query Match 100.0%; Score 380; DB 15; Length 380;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NASLAANAARFCNLFREDDNQNGNVFFSSLSLFAALALVRLGAQDDSLQIDKLLHV 60
DB 1 NASLAANAARFCNLFREDDNQNGNVFFSSLSLFAALALVRLGAQDDSLQIDKLLHV 60
QY 61 NTASGYGNSNSQSGLSQSLKRVFSDINASHKDYDLSIVNGLFAEKVYGFHKDYIECAEK 120
DB 61 NTASGYGNSNSQSGLSQSLKRVFSDINASHKDYDLSIVNGLFAEKVYGFHKDYIECAEK 120
QY 121 LVDKVERVDFTHLEDTRNINKWVENETHGKIKNVIGEGG1SSSAVWLVNAVYFKGK 180
DB 121 LVDKVERVDFTHLEDTRNINKWVENETHGKIKNVIGEGG1SSSAVWLVNAVYFKGK 180
QY 181 WSAFTKSETINCHFKSPKCGKAVAMHQRKFNLSVIEDPSMKILELYNGGINMVL 240
DB 181 WSAFTKSETINCHFKSPKCGKAVAMHQRKFNLSVIEDPSMKILELYNGGINMVL 240
QY 241 LPENDLSIENKLTFOQLMEWTNPRMTSKYVEVFPQFKIEKNYEMKQYLALGLKDI 300
DB 241 LPENDLSIENKLTFOQLMEWTNPRMTSKYVEVFPQFKIEKNYEMKQYLALGLKDI 300
QY 301 DESKADLSGIASGGRLYISRMHKSYLEVTEEGTATAATGTSNIVEKQLPQSTLFRADHP 360
DB 301 DESKADLSGIASGGRLYISRMHKSYLEVTEEGTATAATGTSNIVEKQLPQSTLFRADHP 360
QY 361 FLFVIRKDDIILFSGKVSCP 380
DB 361 FLFVIRKDDIILFSGKVSCP 380
RESULT 2
AAV08254
ID AAV08254 standard; Protein; 380 AA.
AC AAV08254;
XX
XX 14-JUL-1999 (first entry)
XX
XX Human megasin protein.
XX
XX Megsin; mesangial cell; treatment; diagnosis; disease; IgA nephropathy;
KW human; rat; murine.
XX
XX Homo sapiens.
XX
XX W09915652-Al.
XX
XX 01-APR-1999.
XX
XX

```

```

PF 22-SEP-1998; 98WO-JP04269.
XX
PR 22-SEP-1997; 97JP-0275302.
XX
PA (KURO/) KUROKAWA K.
PA (MIYA/) MIYATA T.
XX
PI Miyata T;
XX
XX WPI; 1999-276983/23.
DR N-PSDB; AAX56712.
XX
XX Megsin protein expressed specifically in mesangial cells
PT
PS Claim 1; Page 62-64; 100pp; Japanese.
XX
XX This invention describes the isolation of novel megsin nucleic acid and
XX proteins from human, rat and mouse tissue. This protein is expressed
XX specifically in mesangial cells. The products of the invention are
XX useful for the treatment and diagnosis of diseases involving mesangial
XX cells, such as IgA nephropathy.
XX
SQ Sequence 380 AA;
Query Match 100.0%; Score 380; DB 20; Length 380;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASLAANAARFCNLFREDDNQNGNVFFSSLSLFAALALVRLGAQDDSLQIDKLLHV 60
DB 1 MASLAANAARFCNLFREDDNQNGNVFFSSLSLFAALALVRLGAQDDSLQIDKLLHV 60
QY 61 NTASGYGNSNSQSGLSQSLKRVFSDINASHKDYDLSIVNGLFAEKVYGFHKDYIECAEK 120
DB 61 NTASGYGNSNSQSGLSQSLKRVFSDINASHKDYDLSIVNGLFAEKVYGFHKDYIECAEK 120
QY 121 LVDKVERVDFTHLEDTRNINKWVENETHGKIKNVIGEGG1SSSAVWLVNAVYFKGK 180
DB 121 LVDKVERVDFTHLEDTRNINKWVENETHGKIKNVIGEGG1SSSAVWLVNAVYFKGK 180
QY 181 WSAFTKSETINCHFKSPKCGKAVAMHQRKFNLSVIEDPSMKILELYNGGINMVL 240
DB 181 WSAFTKSETINCHFKSPKCGKAVAMHQRKFNLSVIEDPSMKILELYNGGINMVL 240
QY 241 LPENDLSIENKLTFOQLMEWTNPRMTSKYVEVFPQFKIEKNYEMKQYLALGLKDI 300
DB 241 LPENDLSIENKLTFOQLMEWTNPRMTSKYVEVFPQFKIEKNYEMKQYLALGLKDI 300
QY 301 DESKADLSGIASGGRLYISRMHKSYLEVTEEGTATAATGTSNIVEKQLPQSTLFRADHP 360
DB 301 DESKADLSGIASGGRLYISRMHKSYLEVTEEGTATAATGTSNIVEKQLPQSTLFRADHP 360
QY 361 FLFVIRKDDIILFSGKVSCP 380
DB 361 FLFVIRKDDIILFSGKVSCP 380
RESULT 3
AAB24142
ID AAB24142 standard; Protein; 380 AA.
AC AAB24142;
XX
XX 30-JAN-2001 (first entry)
XX
XX Human megsin protein sequence SEQ ID NO:2.
XX
XX Megsin; mesangium-predominant gene; serpin regulated; nephropathy;
KW IgA; immunoglobulin A; detection; renal function; renal disorder;
KW diagnosis; biological sample; blood; urine.
XX
XX Homo sapiens.
XX

```

FN WC200057189-A1.

XX 28-SEP-2000.

XX 17-MAR-2000; 2000WO-JP01646.

XX 19-MAR-1999; 99JP-0075305.

XX 28-OCT-1999; 99JP-0306623.

XX (KURO/) KUROKAWA K.

XX (FUSO) FUSO PHARM IND LTD.

XX (MIYA/) MIYATA T.

XX Miyata T.

XX WPI; 2000-611642/58.

XX N-PSDB; AAA99294.

XX Evaluating renal function comprises assaying mesgin protein in

XX biological sample -

XX Example 2; Page 66-69; 93pp; Japanese.

XX The present invention describes a method for evaluating renal function.

XX The method comprises assaying mesgin protein in biological sample. Also

XX described are: (1) use of an anti-mesgin protein antibody for diagnosing

XX renal function; and (2) a kit for detecting mesgin protein comprising:

XX (a) anti-mesgin protein antibody attached to solid magnetic particles;

XX (b) direct or indirect fixing for the antibody to the particles; and

XX (c) a magnet. The process is useful for evaluating renal function and

XX diagnosing renal disorders by assaying mesgin protein in biological

XX samples (preferably urine or blood). The process is reproducible and

XX gives accurate results. The present sequence represents the human mesgin

XX protein, which is given in the exemplification of the present invention.

XX Sequence 380 AA;

XX Query Match 100.0%; Score 380; DB 21; Length 380;

XX Best Local Similarity 100.0%; Pred. No. 0;

XX Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX QY 1 MASLAANAECFNLPREDDNQGNGNVFFSSLSLFAALALVRLGAQDLSQIDKLLHV 60

XX Db 1 MASLAANAECFNLPREDDNQGNGNVFFSSLSLFAALALVRLGAQDLSQIDKLLHV 60

XX QY 61 NTASGYGNSNSQSGLSQOLKEVFSIDINASHKDYDLSIVNGLFAEKYGFHXYTECAEK 120

XX Db 61 NTASGYGNSNSQSGLSQOLKEVFSIDINASHKDYDLSIVNGLFAEKYGFHXYTECAEK 120

XX QY 121 LYDAKVERVDFTNHLEDTRNINKWENETHGKIKNVIGEGGSSSAVMVLYNAVYFKGK 180

XX Db 121 LYDAKVERVDFTNHLEDTRNINKWENETHGKIKNVIGEGGSSSAVMVLYNAVYFKGK 180

XX QY 181 WSAFTKSETINCHFKSPKCSKAVAMMHOERKFNLSVIEDPSMKILELYNGGINMYVL 240

XX Db 181 WSAFTKSETINCHFKSPKCSKAVAMMHOERKFNLSVIEDPSMKILELYNGGINMYVL 240

XX QY 241 LPENDLSEINKLTFTONLMEWTNPRMTSKYVEVFPFQFKENYEMKYLALGLKDIF 300

XX Db 241 LPENDLSEINKLTFTONLMEWTNPRMTSKYVEVFPFQFKENYEMKYLALGLKDIF 300

XX QY 301 DESKADLSGIASGGRUYISRMWHKSYIEVTEGTEATAATGSNIVEKQLPQSTLFRADHP 360

XX Db 301 DESKADLSGIASGGRUYISRMWHKSYIEVTEGTEATAATGSNIVEKQLPQSTLFRADHP 360

XX QY 361 FLFVIRKDDIILFSGKVSCP 380

XX Db 361 FLFVIRKDDIILFSGKVSCP 380

XX RESULT 4

XX AAB83075

XX TD AAB83075 standard; Protein; 380 AA.

XX

AC AAB83075;

XX 10-JUL-2001 (first entry)

XX Human mesgin protein.

XX Human; mesgin; mesangial cell proliferative nephritis; nephrotropic;

XX transgenic mouse; glomerular disease; animal model; drug screening.

XX Homo sapiens.

XX WO200124628-A1.

XX 12-APR-2001.

XX 06-OCT-2000; 2000WO-JP06988.

XX 06-OCT-1999; 99JP-0285736.

XX (KURO/) KUROKAWA K.

XX (MIYA/) MIYATA T.

XX Miyata T;

XX WPI; 2001-300136/31.

XX N-PSDB; AAF82438.

XX Mouse model for mesangial cell proliferative nephritis for development

XX and screening of new treatments -

XX Example 4; Page 44-46; 62pp; Japanese.

XX The present sequence is human mesgin. The human mesgin coding

XX sequence may be introduced into a mouse to produce an animal model of

XX mesangial cell proliferative nephritis. The symptoms include

XX enlargement of the mesangial base region, sedimentation of an immune

XX complex and an increase in mesangial cells. The animal model is useful

XX for analysing the pathology of chronic glomerular diseases and for

XX screening compositions for prevention and treatment of the diseases.

XX Highly uniform models can be made easily and in large numbers using

XX this method.

XX Sequence 380 AA;

XX Query Match 100.0%; Score 380; DB 22; Length 380;

XX Best Local Similarity 100.0%; Pred. No. 0;

XX Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX QY 1 MASLAANAECFNLPREDDNQGNGNVFFSSLSLFAALALVRLGAQDLSQIDKLLHV 60

XX Db 1 MASLAANAECFNLPREDDNQGNGNVFFSSLSLFAALALVRLGAQDLSQIDKLLHV 60

XX QY 61 NTASGYGNSNSQSGLSQOLKEVFSIDINASHKDYDLSIVNGLFAEKYGFHXYTECAEK 120

XX Db 61 NTASGYGNSNSQSGLSQOLKEVFSIDINASHKDYDLSIVNGLFAEKYGFHXYTECAEK 120

XX QY 121 LYDAKVERVDFTNHLEDTRNINKWENETHGKIKNVIGEGGSSSAVMVLYNAVYFKGK 180

XX Db 121 LYDAKVERVDFTNHLEDTRNINKWENETHGKIKNVIGEGGSSSAVMVLYNAVYFKGK 180

XX QY 181 WSAFTKSETINCHFKSPKCSKAVAMMHOERKFNLSVIEDPSMKILELYNGGINMYVL 240

XX Db 181 WSAFTKSETINCHFKSPKCSKAVAMMHOERKFNLSVIEDPSMKILELYNGGINMYVL 240

XX QY 241 LPENDLSEINKLTFTONLMEWTNPRMTSKYVEVFPFQFKENYEMKYLALGLKDIF 300

XX Db 241 LPENDLSEINKLTFTONLMEWTNPRMTSKYVEVFPFQFKENYEMKYLALGLKDIF 300

XX QY 301 DESKADLSGIASGGRUYISRMWHKSYIEVTEGTEATAATGSNIVEKQLPQSTLFRADHP 360

XX Db 301 DESKADLSGIASGGRUYISRMWHKSYIEVTEGTEATAATGSNIVEKQLPQSTLFRADHP 360

QY 361 FLFVIRKDDIIILFSGKVSCP 380
 DB 361 FLFVIRKDDIIILFSGKVSCP 380

RESULT 5
 ABUS6533
 ID ABUS6533 standard; Protein; 380 AA.
 XX
 AC ABUS6533;
 XX
 DT 02-APR-2003 (first entry)
 XX
 DE Lung cancer-associated polypeptide #126.
 XX
 KW Lung cancer-associated polypeptide; cytostatic; emphysema;
 KW antinflammatory; anasthmatic; non-small cell lung cancer; atelectasis;
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
 XX
 OS Unidentified.
 XX
 EW WO200256443-A2.
 XX
 PD 31-OCT-2002.
 XX
 PF 16-APR-2002; 2002WO-US12476.
 XX
 PR 18-APR-2001; 2001US-284770P.
 PR 10-MAY-2001; 2001US-290492P.
 PR 09-NOV-2001; 2001US-339245P.
 PR 13-NOV-2001; 2001US-350666P.
 PR 29-NOV-2001; 2001US-334370P.
 PR 12-APR-2002; 2002US-372246P.
 XX
 PA (EOSE-) EOS BIOTECHNOLOGY INC.
 XX
 PI Aziz N, Murray R;
 XX
 DR WFI; 2003-093161/08.
 DR N-PSDB; ABX76260.
 XX
 PT Detecting a lung cancer-associated transcript in a cell from a patient
 PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer -
 XX
 PS Claim 27; Page 288; 453pp; English.
 XX
 CC The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridizes
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung
 CC cancer-associated polynucleotides and polypeptides are used for
 CC identifying a compound that modulates a lung cancer-associated
 CC polypeptide, for inhibiting proliferation of a lung cancer-associated
 CC cell to treat lung cancer in a patient and for treating a mammal having
 CC lung cancer by administering a modulatory compound identified. The
 CC methods are useful for treating lung cancer, such as small cell lung
 CC cancer, non-small cell lung cancer or other benign or precancerous
 CC lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive
 CC pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial
 CC pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides
 CC and polypeptides are useful for diagnostic purposes and as targets for
 CC screening for therapeutic compounds that modulate lung cancer, such as
 CC antibodies. Sequences ABUS6408-ABUS6745 represent lung cancer-associated
 CC polypeptides of the invention.
 XX
 SQ Sequence 380 AA;

Query Match

100.0%; Score 380; DB 24; Length 380;

Best Local Similarity 100.0%; Pred. No. 0;
 Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASLAANAARFCNLFREMDNQGNGVFFSSLSLFAALALVRIGAGQDSSLSQIDKLIHV 60
 DB 1 MASLAANAARFCNLFREMDNQGNGVFFSSLSLFAALALVRIGAGQDSSLSQIDKLIHV 60
 QY 61 NTASGYGNSSNSQSLQSRVFSIDINASHKDYDLSTVNGLFABKVYGFHDKYTECAEK 120
 DB 61 NTASGYGNSSNSQSLQSRVFSIDINASHKDYDLSTVNGLFABKVYGFHDKYTECAEK 120
 QY 121 LYDAKVERVDFTNHLEDTNRNINKWENTHGIKNVIGEGGSSSAVAVLVNAVYFKGK 180
 DB 121 LYDAKVERVDFTNHLEDTNRNINKWENTHGIKNVIGEGGSSSAVAVLVNAVYFKGK 180
 QY 181 WQSAPTKSETINCHPKSPKSGKAVAMHQRKFNLSVIEDPSMKILELFYNGGINMYVL 240
 DB 181 WQSAPTKSETINCHPKSPKSGKAVAMHQRKFNLSVIEDPSMKILELFYNGGINMYVL 240
 QY 241 LPENDLSEIENKLTQNLMEWTNPREMTSKYVEVFFPQFKIEKNYEMKQYLFALGLKDI 300
 DB 241 LPENDLSEIENKLTQNLMEWTNPREMTSKYVEVFFPQFKIEKNYEMKQYLFALGLKDI 300
 QY 301 DESKADLSGIASGRLYTSRMWHSKYIETVETGTEATAATGSGNIVEKQLPQSTILFRADHP 360
 DB 301 DESKADLSGIASGRLYTSRMWHSKYIETVETGTEATAATGSGNIVEKQLPQSTILFRADHP 360
 QY 361 FLFVIRKDDIIILFSGKVSCP 380
 DB 361 FLFVIRKDDIIILFSGKVSCP 380

RESULT 6

AAU08256
 ID AAU08256 standard; Protein; 368 AA.

XX AC AAU08256;
 XX
 DT 14-JUL-1999 (first entry)
 XX
 DE Mouse megasin protein.
 XX
 KW Megasin; mesangial cell; treatment; diagnosis; disease; IgA nephropathy;
 KW human; rat; murine.
 XX
 OS Mus musculus.
 XX
 FH Key
 FT Protein
 FT Location/Qualifiers
 FT 1..368
 FT /note= "partial sequence"
 XX
 FN WO9915652-A1.
 XX
 PD 01-APR-1999.
 XX
 PF 22-SEP-1998; 98WO-JP04269.
 XX
 PR 22-SEP-1997; 97JP-0275302.
 XX
 PA (KURO/) KUROKAWA K.
 PA (MIYA/) MIYATA T.
 XX
 PI Miyata T;
 XX
 DR WFI; 1999-275983/23.
 DR N-PSDB; AAX56714.
 XX
 PT Megasin protein expressed specifically in mesangial cells
 XX
 PS Claim 1; Page 76-79; 100pp; Japanese.

CC This invention describes the isolation of novel megasin nucleic acid and
 CC proteins from human, rat and mouse tissue. This protein is expressed

CC specifically in mesangial cells. The products of the invention are
 CC useful for the treatment and diagnosis of diseases involving mesangial
 CC cells, such as IGA nephropathy.
 XX
 SQ

Sequence 368 AA;

Query Match 4.7%; Score 18; DB 20; Length 368;

Best Local Similarity 100.0%; Pred. No. 2.4e-09;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 SSSAVMLVNAVYFKGW 181

DB 152 SSSAVMLVNAVYFKGW 169

RESULT 7

AAE24151
 ID AAE24151 standard; Protein; 368 AA.

XX AC AAE24151;

XX DT 30-JAN-2001 (first entry)

DE Mouse megin protein sequence SEQ ID NO:21.

XX Megin; mesangium-predominant gene; serpin regulated; nephropathy;
 KW IGA; immunoglobulin A; detection; renal function; renal disorder;
 KW diagnosis; biological sample; blood; urine.

OS Mus musculus.

XX PN WO200057189-A1.

XX PD 28-SEP-2000.

XX PF 17-MAR-2000; 2000WO-JP01646.

XX PR 19-MAR-1999; 99JP-0075305.

XX PS 28-OCT-1999; 99JP-0306623.

XX PA (KURO/) KUROKAWA K.

XX PA (FUSO) FUSO PHARM IND LTD.

XX PA (MIYA/) MIYATA T.

XX PI Miyata T;

XX DR WPI; 2000-611642/56.

XX DR N-PSDB; AAC55239.

XX PT Evaluating renal function comprises assaying megin protein in

XX PT biological sample

XX PS Disclosure; Page 89-91; 93pp; Japanese.

XX The present invention describes a method for evaluating renal function.
 CC The method comprises assaying megin protein in biological sample. Also
 CC described are: (1) use of an anti-megin protein antibody for diagnosing
 CC renal function; and (2) a kit for detecting megin protein comprising:
 CC (a) anti-megin protein antibody attached to solid magnetic particles;
 CC (b) direct or indirect fixing for the antibody to the particles; and
 CC (c) a magnet. The process is useful for evaluating renal function and
 CC diagnosing renal disorders by assaying megin protein in biological
 CC samples (preferably urine or blood). The process is reproducible and
 CC gives accurate results. The present sequence represents the mouse megin
 CC protein, which is given in the exemplification of the present invention.

XX Sequence 368 AA;

Query Match 4.7%; Score 18; DB 21; Length 368;

Best Local Similarity 100.0%; Pred. No. 2.4e-09;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 SSSAVMLVNAVYFKGW 181

DB 152 SSSAVMLVNAVYFKGW 169

RESULT 8

AAE83077

ID AAE83077 standard; Protein; 368 AA.

XX AC AAE83077;

XX DT 10-JUL-2001 (first entry)

XX DE Murine megin protein.

XX XX Mouse; megin; mesangial cell proliferative nephritis; nephrotropic;
 KW transgenic mouse; glomerular disease; animal model; drug screening.

XX OS Mus musculus.

XX PN WO200124628-A1.

XX PD 12-APR-2001.

XX PF 06-OCT-2000; 2000WO-JP06988.

XX PR 06-OCT-1999; 99JP-0285736.

XX PA (KURO/) KUROKAWA K.

XX PA (MIYA/) MIYATA T.

XX PI Miyata T;

XX DR WPI; 2001-300136/31.

XX DR N-PSDB; AAF82440.

XX PT Mouse model for mesangial cell proliferative nephritis for development

XX PS Disclosure; Page 52-53; 62pp; Japanese.

XX The present sequence is murine megin protein. The human megin coding
 CC sequence may be introduced into a mouse to produce an animal model of
 CC mesangial cell proliferative nephritis. The symptoms include
 CC enlargement of the mesangial base region, sedimentation of an immune
 CC complex and an increase in mesangial cells. The animal model is useful
 CC for analysing the pathology of chronic glomerular diseases and for
 CC screening compositions for prevention and treatment of the diseases.
 CC Highly uniform models can be made easily and in large numbers using
 CC this method.

XX Sequence 368 AA;

Query Match 4.7%; Score 18; DB 22; Length 368;

Best Local Similarity 100.0%; Pred. No. 2.4e-09;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 SSSAVMLVNAVYFKGW 181

DB 152 SSSAVMLVNAVYFKGW 169

RESULT 9

AAV08255

ID AAV08255 standard; Protein; 380 AA.

XX AC AAV08255;

XX DT 14-JUL-1999 (first entry)

XX DE Rat megin protein.

XX KW Megin; mesangial cell; treatment; diagnosis; disease; IGA nephropathy;
 KW human; rat; murine.

```

XX OS Rattus rattus.
XX PI WO9915652-A1.
XX PN 01-APR-1999.
XX PD 22-SEP-1998; 98WO-JP04269.
XX PE 22-SEP-1997; 97JP-0275302.
XX PF (KURO/) KUROKAWA K.
XX PG (MIYA/) MIYATA T.
XX PH Miyata T.
XX PI WPI; 1999-276983/23.
XX PJ N-PSDB; AAX56712.
XX PK Megsin protein expressed specifically in mesangial cells
XX PL Claim 1; Page 69-72; 100pp; Japanese.
XX PM This invention describes the isolation of novel megin nucleic acid and
XX PN proteins from human, rat and mouse tissue. This protein is expressed
XX PO specifically in mesangial cells. The products of the invention are
XX PP useful for the treatment and diagnosis of diseases involving mesangial
XX PQ cells, such as IgA nephropathy.
XX PS Sequence 380 AA;
SQ Query Match 4.7%; Score 18; DB 20; Length 380;
Best Local Similarity 100.0%; Pred. No. 2.5e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 SSSAVVVLVNAVYFKGKW 181
DB 164 SSSAVVVLVNAVYFKGKW 181

RESULT 10
AAB24150
ID AAB24150 standard; Protein; 380 AA.
XX AC AAB24150;
XX DT 30-JAN-2001 (first entry)
XX DE Rat megin protein sequence SEQ ID NO:19.
XX KW Megsin; mesangium-predominant gene; serpin regulated; nephropathy;
XX KW IgA; immunoglobulin A; detection; renal function; renal disorder;
XX KW diagnosis; biological sample; blood; urine.
XX OS Rattus norvegicus.
XX PH Key Location/Qualifiers
XX PT Misc-difference 51
XX PT Misc-difference 94 /note= "unspecified"
XX PT Misc-difference 94 /note= "unspecified"
XX PN WO2000057189-A1.
XX PD 28-SEP-2000.
XX PF 17-MAR-2000; 2000WO-JP01646.
XX PG 19-MAR-1999; 95JP-0075305.
XX PH 28-OCT-1999; 95JP-0306623.
XX PI (KURO/) KUROKAWA K.
XX PJ (FUSO ) FUSO PHARM IND LTD.

PA (MIYA/) MIYATA T.
XX PI Miyata T.
XX PN WPI; 2000-611642/58.
XX PD N-PSDB; AAC55238.
XX PE Evaluating renal function comprises assaying megin protein in
XX PF biological sample -
XX PG Example 2; Page 81-84; 93pp; Japanese.
XX PH The present invention describes a method for evaluating renal function.
XX PI The method comprises assaying megin protein in biological sample. Also
XX PJ described are: (1) use of an anti-megsin protein antibody for diagnosing
XX PK renal function; and (2) a kit for detecting megin protein comprising:
XX PL (a) anti-megsin protein antibody attached to solid magnetic particles;
XX PM (b) direct or indirect fixing for the antibody to the particles; and
XX PN (c) a magnet. The process is useful for evaluating renal function and
XX PO diagnosing renal disorders by assaying megin protein in biological
XX PP samples (preferably urine or blood). The process is reproducible and
XX PQ gives accurate results. The present sequence represents the rat megin
XX PS protein, which is given in the exemplification of the present invention.
XX SQ Sequence 380 AA;
Query Match 4.7%; Score 18; DB 21; Length 380;
Best Local Similarity 100.0%; Pred. No. 2.5e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 SSSAVVVLVNAVYFKGKW 181
DB 164 SSSAVVVLVNAVYFKGKW 181

RESULT 11
AAG64286
ID AAG64286 standard; Protein; 380 AA.
XX AC AAG64286;
XX DT 21-SEP-2001 (first entry)
XX DE Rat megin protein.
XX KW Rat; megin; renal mesangial cell; mesangium proliferative nephritis.
XX OS Rattus norvegicus.
XX PN WO200148019-A1.
XX PD 05-JUL-2001.
XX PF 26-DEC-2000; 2000WO-JP09251.
XX PG 28-DEC-1999; 95JP-0373677.
XX PH (KURO/) KUROKAWA K.
XX PI (MIYA/) MIYATA T.
XX PJ Miyata T.
XX PN WPI; 2001-425651/45.
XX PD N-PSDB; AAX48181.
XX PT New antibody recognizing a partial sequence of rat megin protein for
XX PG diagnosis of mesangium proliferative nephritis -
XX PH Disclosure; Page 54-56; 63pp; Japanese.
XX PI The present invention relates to a novel antibody which recognises a
XX PJ peptide consisting of residues 341-354 of rat megin protein. The present
XX PK sequence is the protein sequence for rat megin, which was used in the
XX PL

```


CC present invention. Mesgin is highly expressed in renal mesangial cells
 CC and its level is elevated in mesangium proliferative nephritis. Assay of
 CC the serum or urine level using the antibody is therefore indicative of
 CC this type of disorder.

SX Sequence 380 AA;

Query Match 4.7%; Score 18; DB 22; Length 380;
 Best Local Similarity 100.0%; Pred. No. 2.5e-09;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 SSSAVWLVNVAVFKGW 181
 |||||
 Db 164 SSSAVWLVNVAVFKGW 181

RESULT 12

AA83076
 ID AAB83076 standard; Protein; 380 AA.

AC AAB83076;

DT 10-JUL-2001 (first entry)

DE Rat mesgin protein.

KW Rat; mesgin; mesangial cell proliferative nephritis; nephrotropic;
 KW transgenic mouse; glomerular disease; animal model; drug screening.

OS Rattus norvegicus.

PN WO200124628-A1.

XX 12-APR-2001.

XX 06-OCT-2000; 2000WO-JP06988.

XX 06-OCT-1999; 99JP-0285736.

PA (KURO/) KUROKAWA K.

PA (MIYA/) MIYATA T.

PI Miyata T;

DR WPI: 2001-300136/31.

DR N-PSDB; AAB2439.

XX Mouse model for mesangial cell proliferative nephritis for development
 PT and screening of new treatments -

XX Disclosure; Page 48-50; 62pp; Japanese.

CC The present sequence is rat mesgin. The human mesgin coding
 CC sequence may be introduced into a mouse to produce an animal model of
 CC mesangial cell proliferative nephritis. The symptoms include
 CC enlargement of the mesangial base region, sedimentation of an immune
 CC complex and an increase in mesangial cells. The animal model is useful
 CC for analysing the pathology of chronic glomerular diseases and for
 CC screening compositions for prevention and treatment of the diseases.
 CC Highly uniform models can be made easily and in large numbers using
 CC this method.

SX Sequence 380 AA;

Query Match 4.7%; Score 18; DB 22; Length 380;
 Best Local Similarity 100.0%; Pred. No. 2.5e-09;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 SSSAVWLVNVAVFKGW 181
 |||||
 Db 164 SSSAVWLVNVAVFKGW 181

RESULT 13

AA857112
 ID AAR57112 standard; peptide; 20 AA.

XX AAR57112;

XX 25-MAR-2003 (updated)

DT 16-AUG-1994 (first entry)

XX Human megakaryocyte differentiation factor peptide 6.

KW Human megakaryocyte differentiation factor; MDF; thrombopoietin;
 KW haematopoietic stimulating factor; thrombocytopoiesis; platelet;
 KW bone marrow transplantation; cancer chemotherapy.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 18 /note= "not determined"

XX EP583884-A1.

XX 23-FEB-1994.

XX 19-JUL-1993; 93EP-0305654.

XX 17-JUL-1992; 92JP-0212305.

XX 04-MAR-1993; 93JP-0067339.

XX (SUNR) SUNTORY LTD.

XX (TSUJ) TSUJIMOTO M.

XX Ishida N, Iwasa F, Kurihara T, Miura K, Nakazato H;

PI Tsujimoto M, Tsuruoka N, Yamaguchi N, Yamauchi K;

XX WPI: 1994-058782/08.

XX New megakaryocyte differentiation factor - isolated from human
 PT epidermoid carcinoma cells, used to treat conditions involving a
 PT decrease in platelets

XX Claim 1; Page 20; 47pp; English.

XX Human MDF (see AAR48379) can be isolated from a culture of human
 CC epidermoid carcinoma A431 cells in protein-free medium. The MDF
 CC stimulates differentiation of megakaryocytes from myeloid cells
 CC in the presence of IL-3. The MDF acts in vivo as a thrombopoietin
 CC making it useful for treatment of diseases involving a decrease
 CC in platelet number (esp. thrombocytopaenia) such as occurs in bone
 CC marrow transplantation and in chemotherapy. MDF has mol.wt.
 CC 55-57kD by SDS-PAGE, exhibits an isoelectric point of 6.0-7.0 and
 CC contains an amino acid sequence comprising at least one of the
 CC sequences AAR57107-R57115.
 CC (Updated on 25-MAR-2003 to correct PN field.)

SX Sequence 20 AA;

Query Match 4.5%; Score 17; DB 15; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.5e-09;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 305 ADLSGIASGRLYISRM 321
 |||||

Db 1 ADLSGIASGRLYISRM 17
 |||||

RESULT 14

AA8260
 ID AAY08260 standard; Protein; 16 AA.

XX AAY08260;

DT 14-JUL-1999 (first entry)
 XX
 DE Megin peptide 4.
 XX
 KW Megin; mesangial cell; treatment; diagnosis; disease; IgA nephropathy;
 XX human; rat; murine.
 XX
 OS Unidentified.
 XX
 PN WO9915652-A1.
 XX
 PD 01-APR-1999.
 XX
 PF 22-SEP-1998; 98WO-JF04269.
 XX
 PR 22-SEP-1997; 97JP-0275302.
 XX
 PA (KURO/) KUROKAWA K.
 XX (MIYA/) MIYATA T.
 XX
 PI Miyata T;
 XX
 DR WPI; 1999-276983/23.
 XX
 PT Megin protein expressed specifically in mesangial cells
 XX
 PS Example 10; Page 86; 100pp; Japanese.
 XX
 CC This invention describes the isolation of novel megin nucleic acid and
 CC proteins from human, rat and mouse tissue. This protein is expressed
 CC specifically in mesangial cells. The products of the invention are
 CC useful for the treatment and diagnosis of diseases involving mesangial
 CC cells, such as IgA nephropathy. AA108257-Y08260 represent peptides used
 CC in the method of the invention.
 XX
 SQ Sequence 16 AA;
 Query Match 4.2%; Score 16; DB 20; Length 16;
 Best Local Similarity 100.0%; Pred.No.1.3e-08;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 339 ATGSNIVEKQLPQSTL 354
 DB 1 ATGSNIVEKQLPQSTL 16
 RESULT 15
 AAB24145
 ID AAB24145 standard; Peptide; 16 AA.
 AC AAB24145;
 XX
 DT 30-JAN-2001 (first entry)
 XX
 DE Human megin domain peptide SEQ ID NO:13.
 XX
 KW Megin; mesangium-predominant gene; serpin regulated; nephropathy;
 XX IgA; immunoglobulin A; detection; renal function; renal disorder;
 XX diagnosis; biological sample; blood; urine.
 OS Homo sapiens.
 XX
 PN WO2000057189-A1.
 XX
 PD 28-SEP-2000.
 XX
 PF 17-MAR-2000; 2000WO-JP01646.
 XX
 PR 19-MAR-1999; 99JP-0075305.
 PR 28-OCT-1999; 99JP-0306623.
 XX
 PA (KURO/) KUROKAWA K.
 XX (FUSO) FUSO PHARM IND LTD.

PA (MIYA/) MIYATA T.
 XX
 PI Miyata T;
 XX
 DR WPI; 2000-611642/58.
 XX
 PT Evaluating renal function comprises assaying megin protein in
 PT biological sample -
 XX
 PS Example 2; Page 29; 93pp; Japanese.
 XX
 CC The present invention describes a method for evaluating renal function.
 CC The method comprises assaying megin protein in biological sample. Also
 CC described are: (1) use of an antimegin protein antibody for diagnosing
 CC renal function; and (2) a kit for detecting megin protein comprising:
 CC (a) antimegin protein antibody attached to solid magnetic particles;
 CC (b) direct or indirect fixing for the antibody to the particles; and
 CC (c) a magnet. The process is useful for evaluating renal function and
 CC diagnosing renal disorders by assaying megin protein in biological
 CC samples (preferably urine or blood). The process is reproducible and
 CC gives accurate results. The present sequence represents a human megin
 CC domain peptide, which is used in an example from the present invention.
 XX
 SQ Sequence 16 AA;
 Query Match 4.2%; Score 16; DB 21; Length 16;
 Best Local Similarity 100.0%; Pred.No.1.3e-08;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 339 ATGSNIVEKQLPQSTL 354
 DB 1 ATGSNIVEKQLPQSTL 16
 Search completed: December 12, 2003, 16:33:18
 Job time : 44 secs

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 4-212305
;; FILING DATE: 17-JUL-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 6-067339
;; FILING DATE: 04-MAR-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: McGowan, Malcolm K.
;; REGISTRATION NUMBER: 39,300
;; REFERENCE/DOCKET NUMBER: 001560-247
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 836-6620
;; TELEFAX: (703) 836-2021
;; INFORMATION FOR SEQ ID NO: 34:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 380 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-09-140-719-34

Query Match 100.0%; Score 380; DB 9; Length 380;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MASLAANAEEFCNLFREMDNQGNGVFFSSLSFALALVRLGQDDLSQIDKLLHV 60
Db 1 MASLAANAEEFCNLFREMDNQGNGVFFSSLSFALALVRLGQDDLSQIDKLLHV 60
Qy 61 NTASGYGNSNSQSGLSQQLKRVFSDINASHKDYDLSVINGLFAEKVYGFHKDYIECAEK 120
Db 61 NTASGYGNSNSQSGLSQQLKRVFSDINASHKDYDLSVINGLFAEKVYGFHKDYIECAEK 120
Qy 121 LYDAKVRVDFTHLEDTNRNINKWVENETHGKIKNVIGEGGISSAVMVLNVAVYFKGK 180
Db 121 LYDAKVRVDFTHLEDTNRNINKWVENETHGKIKNVIGEGGISSAVMVLNVAVYFKGK 180
Qy 181 WSAFTKSETINCHFKSPKCSKAVAMHQRKFNLSVIEDPSMKILELRVNGGINVYL 240
Db 181 WSAFTKSETINCHFKSPKCSKAVAMHQRKFNLSVIEDPSMKILELRVNGGINVYL 240
Qy 241 LPENDLSEIENKLTFTONLWENTNPRMTSKYVEVFFPFQFKIEKNYEMKQYLRALGLKDI 300
Db 241 LPENDLSEIENKLTFTONLWENTNPRMTSKYVEVFFPFQFKIEKNYEMKQYLRALGLKDI 300
Qy 301 DESKADLSGIASGRLYISRMMHKSYLEVTEGTATATGTSNIVEKQLPQSTLFRADHP 360
Db 301 DESKADLSGIASGRLYISRMMHKSYLEVTEGTATATGTSNIVEKQLPQSTLFRADHP 360
Qy 361 FLFVIRKDDIILFSGKVSCP 380
Db 361 FLFVIRKDDIILFSGKVSCP 380

RESULT 2

US-10-091-442-34
; Sequence 34, Application US/10091442
; Publication No. US20020164711A1
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; TWASA, Fuyuki
; TSURUOKA, No. US20020164711A1
; NAKAZATO, Hiroshi
; MIURA, Kenju
; ISHIDA, No. US20020164711A1
; KURIHARA, Tatsuya
; YAMAUCHI, Kozo
; YAMAGUCHI, No. US20020164711A1
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404

;; CITY: Alexandria
;; STATE: Virginia
;; COUNTRY: United States
;; ZIP: 22313-1404
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent in Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/10/091,442
;; FILING DATE: 07-MAR-2002
;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/09/140,719
;; FILING DATE: 08-AUG-1998
;; APPLICATION NUMBER: US 08/474,661
;; FILING DATE: 07-JUN-1995
;; APPLICATION NUMBER: US 08/091,028
;; FILING DATE: 14-JUL-1993
;; APPLICATION NUMBER: JP 4-212305
;; FILING DATE: 17-JUL-1992
;; APPLICATION NUMBER: JP 6-067339
;; FILING DATE: 04-MAR-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: McGowan, Malcolm K.
;; REGISTRATION NUMBER: 39,300
;; REFERENCE/DOCKET NUMBER: 001560-247
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 836-6620
;; TELEFAX: (703) 836-2021
;; INFORMATION FOR SEQ ID NO: 34:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 380 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-10-091-442-34

Query Match 100.0%; Score 380; DB 14; Length 380;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MASLAANAEEFCNLFREMDNQGNGVFFSSLSFALALVRLGQDDLSQIDKLLHV 60
Db 1 MASLAANAEEFCNLFREMDNQGNGVFFSSLSFALALVRLGQDDLSQIDKLLHV 60
Qy 61 NTASGYGNSNSQSGLSQQLKRVFSDINASHKDYDLSVINGLFAEKVYGFHKDYIECAEK 120
Db 61 NTASGYGNSNSQSGLSQQLKRVFSDINASHKDYDLSVINGLFAEKVYGFHKDYIECAEK 120
Qy 121 LYDAKVRVDFTHLEDTNRNINKWVENETHGKIKNVIGEGGISSAVMVLNVAVYFKGK 180
Db 121 LYDAKVRVDFTHLEDTNRNINKWVENETHGKIKNVIGEGGISSAVMVLNVAVYFKGK 180
Qy 181 WSAFTKSETINCHFKSPKCSKAVAMHQRKFNLSVIEDPSMKILELRVNGGINVYL 240
Db 181 WSAFTKSETINCHFKSPKCSKAVAMHQRKFNLSVIEDPSMKILELRVNGGINVYL 240
Qy 241 LPENDLSEIENKLTFTONLWENTNPRMTSKYVEVFFPFQFKIEKNYEMKQYLRALGLKDI 300
Db 241 LPENDLSEIENKLTFTONLWENTNPRMTSKYVEVFFPFQFKIEKNYEMKQYLRALGLKDI 300
Qy 301 DESKADLSGIASGRLYISRMMHKSYLEVTEGTATATGTSNIVEKQLPQSTLFRADHP 360
Db 301 DESKADLSGIASGRLYISRMMHKSYLEVTEGTATATGTSNIVEKQLPQSTLFRADHP 360
Qy 361 FLFVIRKDDIILFSGKVSCP 380
Db 361 FLFVIRKDDIILFSGKVSCP 380

RESULT 3

US-09-140-719-3
; Sequence 3, Application US/09140719
; Patent No. US20010026931A1
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Fuyuki
; APPLICANT: TSURUOKA, No. US20010026931A1uo
; APPLICANT: NAKAZATO, Hiroshi
; APPLICANT: MURA, Kenju
; APPLICANT: ISHIDA, No. US20010026931A1uhiro
; APPLICANT: KURIHARA, Tatsuya
; APPLICANT: YAMAICHI, Kozo
; APPLICANT: YAMAGUCHI, No. US20010026931A1omi
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/140,719
; FILING DATE: 08-AUG-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/474,661
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-067339
; FILING DATE: 04-WAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-247
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-140-719-3

Query Match 5.0%; Score 19; DB 9; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.5e-11;
Matches 19; Conservative 0; Mismatches 0; Indels 0;

QY 126 VERVDFTNHLEDFRRNINK 144
DB 1 VERVDFTNHLEDFRRNINK 19

RESULT 4

US-10-091-442-3
; Sequence 3, Application US/10091442
; Publication No. US20020164711A1

GENERAL INFORMATION:

APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURUOKA, No. US20020164711A1uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MURA, Kenju
APPLICANT: ISHIDA, No. US20020164711A1uhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. US20020164711A1omi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/091,442
FILING DATE: 07-Mar-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-WAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-091-442-3

Query Match 5.0%; Score 19; DB 14; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.5e-11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 126 VERVDFTNHLEDFRRNINK 144
DB 1 VERVDFTNHLEDFRRNINK 19

RESULT 5

US-09-140-719-6
; Sequence 6, Application US/09140719
; Patent No. US20010026931A1
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Fuyuki
; APPLICANT: TSURUOKA, No. US20010026931A1uo

APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. US20010026931A1uhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. US20010026931A1omi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Peptide
US-09-140-719-6

Query Match 4.7%; Score 18; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.6e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 305 ADLSGIASGGRLYISRM 322
Db 1 ADLSGIASGGRLYISRM 18

RESULT 6
US-10-091-442-6
Sequence 6, Application US/10091442
Publication No. US20020164711A1
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURUOKA, No. US20020164711A1u
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. US20020164711A1uhiro
APPLICANT: KURIHARA, Tatsuya

YAMAICHI, Kozo
YAMAGUCHI, No. US20020164711A1omi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/091,442
FILING DATE: 07-Mar-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-091-442-6

Query Match 4.7%; Score 18; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.6e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 305 ADLSGIASGGRLYISRM 322
Db 1 ADLSGIASGGRLYISRM 18

RESULT 7
US-09-140-719-4
Sequence 4, Application US/09140719
Patent No. US20010026931A1
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURUOKA, No. US20010026931A1u
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. US20010026931A1uhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. US20010026931A1omi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Burns, Doane, Swecker & Mathis
;; STREET: P.O. Box 1404
;; CITY: Alexandria
;; STATE: Virginia
;; COUNTRY: United States
;; ZIP: 22313-1404
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/140,719
;; FILING DATE: 08-AUG-1998
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/474,661
;; FILING DATE: 07-JUN-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/091,028
;; FILING DATE: 14-JUL-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 4-212305
;; FILING DATE: 17-JUL-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 6-067339
;; FILING DATE: 04-MAR-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: McGowan, Malcolm K.
;; REGISTRATION NUMBER: 39,300
;; REFERENCE/DOCKET NUMBER: 001560-247
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 836-6620
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 17 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; US-09-140-719-4

Query Match 4.5%; Score 17; DB 9; Length 17;
Best Local Similarity 100.0%; Pred. No. 2.2e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 325 SYIEVTEGTEATAATG 341
Db 1 SYIEVTEGTEATAATG 17

RESULT 8
US-10-091-442-4
; Sequence 4, Application US/10091442
; Publication No. US20020164711a1
; GENERAL INFORMATION:
; APPLICANT: TSUCIMOTO, Masafumi
; TWASA, Fuyuki
; TSURUOKA, No. US20020164711a1uo
; NAKAZATO, Hiroshi
; MIURA, Kenju
; ISHIDA, No. US20020164711a1uhiro
; KORIHARA, Tatsuya
; YAMAICHI, Kozo
; YAMAGUCHI, No. US20020164711a1oni
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria

;; STATE: Virginia
;; COUNTRY: United States
;; ZIP: 22313-1404
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/10/091,442
;; FILING DATE: 07-Mar-2002
;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/09/140,719
;; FILING DATE: 08-AUG-1998
;; APPLICATION NUMBER: US 08/474,661
;; FILING DATE: 07-JUN-1995
;; APPLICATION NUMBER: US 08/091,028
;; FILING DATE: 14-JUL-1993
;; APPLICATION NUMBER: JP 4-212305
;; FILING DATE: 17-JUL-1992
;; APPLICATION NUMBER: JP 6-067339
;; FILING DATE: 04-MAR-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: McGowan, Malcolm K.
;; REGISTRATION NUMBER: 39,300
;; REFERENCE/DOCKET NUMBER: 001560-247
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 836-6620
;; TELEFAX: (703) 836-2021
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 17 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-091-442-4

Query Match 4.5%; Score 17; DB 14; Length 17;
Best Local Similarity 100.0%; Pred. No. 2.2e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 325 SYIEVTEGTEATAATG 341
Db 1 SYIEVTEGTEATAATG 17

RESULT 9
US-09-902-684-4
; Sequence 4, Application US/09902684
; Patent No. US20020127840A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: PANCREAS-DERIVED PLASMINOGEN ACTIVATOR
; INHIBITOR
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/902,684
; FILING DATE: 12-Jul-2001

```
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 09/026,408
/ FILING DATE: <Unknown>
/ APPLICATION NUMBER: US 60/024,056
/ FILING DATE: 16-AUG-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: STEFFE, ERIC K.
/ REGISTRATION NUMBER: 36,688
/ REFERENCE/DOCKET NUMBER: 1488.0300002
/ TELEPHONE: 202-371-2600
/ TELEFAX: 202-371-2540
/ INFORMATION FOR SEQ ID NO: 4:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 415 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-902-684-4

Query Match 3.4%; Score 13; DB 10; Length 415;
Best Local Similarity 100.0%; Pred. No. 0.00035;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 MVLNNAVYFKGW 181
DB 196 MVLNNAVYFKGW 208
|||||

RESULT 10
US-10-094-944-10
/ Sequence 10, Application US/10094944
/ Publication No. US20030180275A1
/ GENERAL INFORMATION:
/ APPLICANT: Immunex Corporation
/ APPLICANT: Clarke, Howard RG
/ APPLICANT: Dioso, Robert F
/ APPLICANT: Wiley, Steven R
/ TITLE OF INVENTION: HUMAN SERPIN POLYPEPTIDES
/ FILE REFERENCE: 3223-A
/ CURRENT APPLICATION NUMBER: US/10/094,944
/ CURRENT FILING DATE: 2002-03-08
/ NUMBER OF SEQ ID NOS: 35
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 10
/ LENGTH: 415
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-094-944-10

Query Match 3.4%; Score 13; DB 12; Length 415;
Best Local Similarity 100.0%; Pred. No. 0.00035;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 MVLNNAVYFKGW 181
DB 196 MVLNNAVYFKGW 208
|||||

RESULT 11
US-10-017-129-4
/ Sequence 4, Application US/10017128
/ Publication No. US20030124536A1
/ GENERAL INFORMATION:
/ APPLICANT: McCarthy, Jeanette
/ TITLE OF INVENTION: DIAGNOSIS AND TREATMENT FOR VASCULAR DISEASE
/ FILE REFERENCE: MM-001
/ CURRENT APPLICATION NUMBER: US/10/017,128
/ CURRENT FILING DATE: 2001-12-14
/ PRIOR APPLICATION NUMBER: US 60/306,941

/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 09/026,408
/ FILING DATE: <Unknown>
/ APPLICATION NUMBER: US 60/024,056
/ FILING DATE: 16-AUG-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: STEFFE, ERIC K.
/ REGISTRATION NUMBER: 36,688
/ REFERENCE/DOCKET NUMBER: 1488.0300002
/ TELEPHONE: 202-371-2600
/ TELEFAX: 202-371-2540
/ INFORMATION FOR SEQ ID NO: 4:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 415 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-902-684-4

Query Match 3.4%; Score 13; DB 10; Length 415;
Best Local Similarity 100.0%; Pred. No. 0.00035;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 MVLNNAVYFKGW 181
DB 196 MVLNNAVYFKGW 208
|||||

RESULT 12
US-10-106-698-4538
/ Sequence 4538, Application US/10106698
/ Publication No. US20030109690A1
/ GENERAL INFORMATION:
/ APPLICANT: Ruben et al.
/ TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypept
/ FILE REFERENCE: PA005P1
/ CURRENT APPLICATION NUMBER: US/10/106,698
/ CURRENT FILING DATE: 2002-03-27
/ PRIOR APPLICATION NUMBER: PCT/US00/26524
/ PRIOR FILING DATE: 2000-09-28
/ PRIOR APPLICATION NUMBER: US 60/157,137
/ PRIOR FILING DATE: 1999-09-29
/ PRIOR APPLICATION NUMBER: US 60/163,280
/ PRIOR FILING DATE: 1999-11-03
/ NUMBER OF SEQ ID NOS: 8564
/ SOFTWARE: Patent Ver. 3.0
/ SEQ ID NO 4538
/ LENGTH: 423
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-106-698-4538

Query Match 3.4%; Score 13; DB 15; Length 423;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 MVLNNAVYFKGW 181
DB 204 MVLNNAVYFKGW 216
|||||

RESULT 13
US-09-140-719-9
/ Sequence 9, Application US/09140719
/ Patent No. US20010026931A1
/ GENERAL INFORMATION:
/ APPLICANT: TSUJIMOTO, Masafumi
/ APPLICANT: IWASA, Fuyuki
/ APPLICANT: TSUBOYOKA, No. US20010026931A1uo
/ APPLICANT: NAKAZATO, Hiroshi
/ APPLICANT: MIURA, Kenju
/ APPLICANT: ISHIDA, No. US20010026931A1uhiro
/ APPLICANT: KURIHARA, Tatsuya
/ APPLICANT: YAMAICHI, Kozi
/ APPLICANT: YAMAGUCHI, No. US20010026931A1omi
/ TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
/ NUMBER OF SEQUENCES: 34
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Burns, Doane, Swecker & Mathis
```



```

; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/140,719
; FILING DATE: 08-AUG-1998
; CLASSIFICATION:
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/474,661
; FILING DATE: 07-JUN-1995
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-067339
; FILING DATE: 04-MAR-1993
;
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-247
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
;
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
;
; US-09-140-719-9
;
; Query Match 2.6%; Score 10; DB 9; Length 10;
; Best Local Similarity 100.0%; Pred. No. 0.011;
; Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
QY 204 AVAMMHQERK 213
Db 1 AVAMMHQERK 10
;
; RESULT 14
; US-10-091-442-9
; Sequence 9, Application US/10091442
; Publication No. US20020164711A1
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; INASA, Fuyuki
; TSURUOKA, No. US20020164711A1uo
; NAKAZATO, Hiroshi
; MIURA, Kenju
; ISHIDA, No. US20020164711A1uhiro
; KURIHARA, Tatsuya
; YAMAICHI, Koza
; YAMAGUCHI, No. US20020164711A1omi
; TITLE OF INVENTION: MGCARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States

```

```

; ZIP: 22313-1404
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/091,442
; FILING DATE: 07-Mar-2002
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/140,719
; FILING DATE: 08-AUG-1998
;
; APPLICATION NUMBER: US 08/474,661
; FILING DATE: 07-JUN-1995
;
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
;
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
;
; APPLICATION NUMBER: JP 6-067339
; FILING DATE: 04-MAR-1993
;
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-247
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
;
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 9:
;
; US-10-091-442-9
;
; Query Match 2.6%; Score 10; DB 14; Length 10;
; Best Local Similarity 100.0%; Pred. No. 0.011;
; Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
QY 204 AVAMMHQERK 213
Db 1 AVAMMHQERK 10
;
; RESULT 15
; US-09-993-180-45
; Sequence 45, Application US/09993180
; Publication No. US20030054445A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN SERPIN SECRETED FROM LY
; TITLE OF INVENTION: CELLS, LSI-01
; FILE REFERENCE: D0051.NP
; CURRENT APPLICATION NUMBER: US/09/993,180
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 60/248,434
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: US 60/257,610
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: US 60/282,745
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 45
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-993-180-45
;
; Query Match 2.6%; Score 10; DB 11; Length 16;

```

Best Local Similarity 100.0%; Pred. No. 0.017;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 331 EGGTATAAT 340
|||||||
Db 4 EGGTATAAT 13

Search completed: December 12, 2003, 16:32:21
Job time : 31 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 16:28:03 ; Search time 17 Seconds
(without alignments)
1051.185 Million cell updates/sec

Title: US-09-508-997A-2

Perfect score: 380

Sequence: 1 MASLAANAFSCNLFREMD.....FLFVIRKDDIILFSGKVSCP 380

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	380	100.0	380	1	SPB7_HUMAN
2	18	4.7	380	1	Q96595 mus musculus
3	13	3.4	415	1	PA12_HUMAN
4	13	3.4	415	1	PA12_MOUSE
5	10	2.6	397	1	GDN_MOUSE
6	10	2.6	397	1	GDN_RAT
7	10	2.6	398	1	GDN_HUMAN
8	9	2.4	214	1	SI21_RAT
9	9	2.4	232	1	OVAX_CHICK
10	9	2.4	378	1	PT16_BOVIN
11	9	2.4	382	1	OTVL_COTJA
12	9	2.4	385	1	OVAX_CHICK
13	9	2.4	388	1	OVAX_CHICK
14	9	2.4	405	1	SB12_HUMAN
15	9	2.4	413	1	CPT3_RAT
16	8	2.1	291	1	CD20_MOUSE
17	8	2.1	376	1	PT16_HUMAN
18	8	2.1	391	1	SB13_HUMAN
19	8	2.1	423	1	AACT_HUMAN
20	8	2.1	501	1	YSV1_CAEEL
21	8	2.1	922	1	B3AT_CHICK
22	7	1.8	91	1	MERP_ALCSP
23	7	1.8	91	1	MERP_SALTI
24	7	1.8	100	1	VAM3_HUMAN
25	7	1.8	110	1	CHHL_BOMBO
26	7	1.8	165	1	P2Y4_CRIGR
27	7	1.8	170	1	HPAC_SALDU
28	7	1.8	170	1	HPAC_SALTI
29	7	1.8	170	1	HPAC_SALTY
30	7	1.8	188	1	AX2C_PHAUF
31	7	1.8	240	1	YCXN_PORFJ
32	7	1.8	297	1	CD20_HUMAN
33	7	1.8	324	1	DLH1_CANAL

34	7	1.8	361	1	P2Y4_RAT
35	7	1.8	372	1	YD80_BACHD
36	7	1.8	374	1	SPB8_HUMAN
37	7	1.8	375	1	MASP_RAT
38	7	1.8	376	1	SPB9_HUMAN
39	7	1.8	378	1	PT16_MOUSE
40	7	1.8	379	1	ILEU_HUMAN
41	7	1.8	382	1	NDF2_HUMAN
42	7	1.8	392	1	SB11_HUMAN
43	7	1.8	396	1	AFS1_AQUAR
44	7	1.8	410	1	NEUS_CHICK
45	7	1.8	410	1	NEUS_HUMAN

ALIGNMENTS

RESULT 1
SPB7_HUMAN
ID SPB7_HUMAN STANDARD; PRT; 380 AA.
AC 075635;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Megsin (TF55) (Serpin B7).
GN SERPINE7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97326116; PubMed=9182567;
RA Tsujimoto K., Tsuruoka N., Ishida N., Kurihara T., Iwasa F.,
Yamashiro K., Eggi T., Kodama S., Katsuragi N., Adachi M.,
Katayama T., Nakao M., Yamashiro K., Hashino J., Haruyama M., Miura K.,
Nakanishi T., Nakazato H., Teramura M., Mizoguchi H., Yamaguchi S.,
RT "Purification, cDNA cloning, and characterization of a new serpin with
RT megakaryocyte maturation activity".
RL J. Biol. Chem. 272:15373-15380(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Mesangial cells;
RA MEDLINE=98376492; PubMed=9710452;
Miyata T., Nargaku M., Suzuki D., Inagi R., Uragami K., Sakai H.,
Okubo K., Kurokawa K.;
RT "A mesangium-predominant gene, megsin, is a new serpin upregulated in
RT IgA nephropathy".
RL Clin. Invest. 102:828-836(1998).
CC -!- FUNCTION: Might function as an inhibitor of lys-specific
proteases. Might influence the maturation of megakaryocytes via
its action as a serpin.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- TISSUE SPECIFICITY: Predominantly expressed in mesangial cells.
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D88575; BAA31232.1; -.
CC EMBL; AF027866; AAC64506.1; -.
CC HSP; P05619; IHLB.
CC Genew; HGNC:13902; SERPINE7.
CC MIM; 603357; -.
CC GO; GO:0004868; F:serpin; TAS.
CC InterPro; IPR000215; Serpin.
CC Pfam; PF00079; serpin; 1.
CC SMART; SM00063; SERPIN; 1.

```

DR PROSITE; PS00284; SERPIN; 1.
KW Serpin: Serine protease inhibitor.
FT ACT SITE 347 348 REACTIVE BOND (BY SIMILARITY).
SQ SEQUENCE 380 AA; 42904 MW; 9A2CDB6C63CFF605 CRC64;

Query Match 100.0%; Score 380; DB 1; Length 380;
Best Local Similarity 100.0%; Pred. No. 0; Indels 0; Gaps 0;
Matches 380; Conservative 0; Mismatches 0;

QY 1 MASLAANAARECFNLFFEMDDNGNGNFFSSLSLFAALALVRLGAQDLSIQDKLLHV 60
DB 1 MASLAANAARECFNLFFEMDDNGNGNFFSSLSLFAALALVRLGAQDLSIQDKLLHV 60
QY 61 NTASGCGSSNSQSGLOSLKRVFSDINASHKDYDLSVNGLFARKYVGHKDYIECAEK 120
DB 61 NTASGCGSSNSQSGLOSLKRVFSDINASHKDYDLSVNGLFARKYVGHKDYIECAEK 120
QY 121 LYDAKVERVDFTNHLETRNINKWVENETGKIKVIGEGGSISSAVWLVNAVYFKGK 180
DB 121 LYDAKVERVDFTNHLETRNINKWVENETGKIKVIGEGGSISSAVWLVNAVYFKGK 180
QY 181 WQSAFTKGETTINCHFKSPKCGKAVAMQERKENLSVIEDPSMKILELYNGINMYVL 240
DB 181 WQSAFTKGETTINCHFKSPKCGKAVAMQERKENLSVIEDPSMKILELYNGINMYVL 240
QY 241 LPENDLSIENKLTFOQLMEWTNPRMTSKYVEVFPQPKIEKNYEMKQYLRALGKQIF 300
DB 241 LPENDLSIENKLTFOQLMEWTNPRMTSKYVEVFPQPKIEKNYEMKQYLRALGKQIF 300
QY 301 DESKADLSIASGGLRYSIMMKXSVIEVTEGTEATATGTSNIVEKQLFQSTLFRADHP 360
DB 301 DESKADLSIASGGLRYSIMMKXSVIEVTEGTEATATGTSNIVEKQLFQSTLFRADHP 360
QY 361 FLFVRKDDIILFGSKVSCP 380
DB 361 FLFVRKDDIILFGSKVSCP 380

RESULT 2
SPB7_MOUSE STANDARD; PRT; 380 AA.
AC Q9D695;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Messin (Serpin B7).
GN SERPINB7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RX MEDLINE=21368006; PubMed=11473647;
RA Nangaku M., Miyata T., Suzuki D., Umezono T., Hashimoto T., Wada T.,
RA Yagi M., Nagao N., Inagi R., Kurokawa K.;
RT "Cloning of rodent megin revealed its up-regulation in
RT mesangio proliferative nephritis.";
RL Kidney Int. 60:641-652(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann N., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

```

```

Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
Norgone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
Suzuki H., Toyooka K., Wang K.H., Weltz C., Whittaker C., Wilming L.,
Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
Hayashizaki Y.;
R "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC 1- FUNCTION: Might function as an inhibitor of Lys-specific
CC proteases. Might influence the maturation of megakaryocytes via
CC its action as a serpin (By similarity).
CC 2- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC 3- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
CC 4- This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC 5-
CC EMBL; AP105328; AAL16768.1; -
CC EMBL; AK014524; BAB29410.1; -
CC HSSP; P05121; 1A7C.
CC MGD; MGI:2151053; Serpinb7.
CC InterPro; IPR000215; Serpinb7.
CC Pfam; PF00079; serpin; 1.
CC SMART; SM00093; SERPIN; 1.
CC PROSITE; PS00284; SERPIN; 1.
KW Serpin; Serine protease inhibitor.
FT ACT SITE 347 348 REACTIVE BOND (BY SIMILARITY).
SQ SEQUENCE 380 AA; 43050 MW; C9240272ECF9CF4 CRC64;

Query Match 4.7%; Score 18; DB 1; Length 380;
Best Local Similarity 100.0%; Pred. No. 1; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 0;

QY 164 SSSAVWLVNAVYFKGW 181
DB 164 SSSAVWLVNAVYFKGW 181

RESULT 3
PAI2_HUMAN STANDARD; PRT; 415 AA.
ID PAI2_HUMAN STANDARD; PRT; 415 AA.
AC P05120; 096E96;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Plasminogen activator inhibitor-2 precursor (PAI-2) (Placental
DE plasminogen activator inhibitor) (Monocyte Arg-serpin) (Urokinase
DE inhibitor).
GN SERPIN2 OR PAI2 OR PLANH2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=87137674; PubMed=3029122;
RA Ye R.D., Wu T.-Z., Sadler J.E.;
RT "cDNA cloning and expression in Escherichia coli of a plasminogen
RT activator inhibitor from human placenta.";
RL J. Biol. Chem. 262:3718-3725(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88142852; PubMed=3325828;
RA Schleuning W.-D., Medcalf R.L., Hession C., Rothenbuhler R.,
RA Shaw A., Kruithof E.K.O.;

```

RT "Plasminogen activator inhibitor 2: regulation of gene transcription
 RT during phorbol ester-mediated differentiation of U-937 human
 RT histiocytic lymphoma cells.";
 RL Mol. Cell. Biol. 7:4564-4567(1987).
 RN [3]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=89174589; PubMed=2494165;
 RX Ye R.D., Ahern S.M., le Beau M.M., Lebo R.V., Sadler J.B.;
 RA "Structure of the gene for human plasminogen activator inhibitor-2.
 RT The nearest mammalian homologue of chicken ovalbumin.";
 RL J. Biol. Chem. 264:5495-5502(1989).
 RN [4]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Monocytes;
 RX MEDLINE=87252928; PubMed=3496414;
 RA Webb A.C., Collins K.L., Snyder S.F., Alexander S.J.,
 RA Rosenwasser L.U., Eddy R.L., Shows T.B., Auron P.E.;
 RT "Human monocyte Arg-Serpin cDNA. Sequence, chromosomal assignment,
 RT and homology to plasminogen activator-inhibitor.";
 RL J. Exp. Med. 166:77-94(1987).
 RN [5]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Monocytes;
 RX MEDLINE=48125032; PubMed=3257578;
 RA Antalis T.M., Clark M.A., Barres T., Lehrbach P.P., Devine P.L.,
 RA Schevov G., Goss N.H., Stephens R.W., Tolstoshev P.;
 RT "Cloning and expression of a cDNA coding for a human monocyte-derived
 RT plasminogen activator inhibitor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:985-989(1988).
 RN [6]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=90152678; PubMed=2303256;
 RX Samia J.A., Alexander S.J., Horton K.W., Auron P.E., Byers M.G.,
 RA Shows T.B. Jr., Webb A.C.;
 RT "Chromosomal organization and localization of the human urokinase
 RT inhibitor gene: perfect structural conservation with ovalbumin.";
 RL Genomics 6:159-167(1990).
 RN [7]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Brain;
 RX MEDLINE=22389257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenon C.M., Schuler G.D.,
 RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Toshiyuki S., Carninci P., Prange C.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Abramson R.D., Mullaly S.J.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Scallus D.E.,
 RA Scinnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16099-16903(2002).
 RN [8]
 RN SEQUENCE OF 12-17; 103-108 AND 314-321.
 RX MEDLINE=93162043; PubMed=1296667;
 RA Rasmussen E.H., van Damme U., Puype M., Gesser B., Celis J.E.,
 RA Vandekerckhove J.;
 RT "Microsequences of 145 proteins recorded in the two-dimensional gel
 RT protein database of normal human epidermal keratinocytes.";
 RL Electrophoresis 13:960-969(1992).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RX MEDLINE=99148112; PubMed=10368272;
 RA Harrop S.J., Jankova L., Coles M., Jardine D., Whittaker J.S.,

RA Gould A.R., Weister A., King G.C., Mabbutt B.C., Curmi P.M.G.;
 RT "The crystal structure of plasminogen activator inhibitor 2 at 2.0-A
 RT resolution: implications for serpin function.";
 RL Structure 7:43-54(1999).
 RN [10]
 RN VARIANTS ASP-120; HIS-229; LYS-404 AND CYS-413.
 RX MEDLINE=99318093; PubMed=10391209;
 RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemes J., Ziaugra L.,
 RA Friedland L., Rolfe A., Warrington J., Lipschutz R., Daley G.Q.,
 RA Lander E.S.;
 RT "Characterization of single-nucleotide polymorphisms in coding regions
 RT of human genes.";
 RL Nat. Genet. 22:231-238(1999).
 RN [11]
 RN ERRATUM.
 RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemes J., Ziaugra L.,
 RA Friedland L., Rolfe A., Warrington J., Lipschutz R., Daley G.Q.,
 RA Lander E.S.;
 RL Nat. Genet. 23:373-373(1999).
 CC -1- FUNCTION: PAI-2 INHIBITS UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
 CC THE MONOCYTE DERIVED PAI-2 IS DISTINCT FROM THE ENDOTHELIAL
 CC CELL-DERIVED PAI-1.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC OR EXTRACELLULAR.
 CC -1- PTM: THE SIGNAL SEQUENCE IS NOT CLEAVED.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use, by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; J02685; AAA36413.1; -;
 DR EMBL; M18082; AAA60006.1; -;
 DR EMBL; Y00630; CAA68666.1; -;
 DR EMBL; M24557; AAA60348.1; -;
 DR EMBL; M24551; AAA60348.1; JOINED.
 DR EMBL; M24552; AAA60348.1; JOINED.
 DR EMBL; M24653; AAA60348.1; JOINED.
 DR EMBL; M24654; AAA60348.1; JOINED.
 DR EMBL; M24655; AAA60348.1; JOINED.
 DR EMBL; M24656; AAA60348.1; JOINED.
 DR EMBL; J03603; AAA60004.1; -;
 DR EMBL; M31551; AAA36797.1; -;
 DR EMBL; M31547; AAA36797.1; JOINED.
 DR EMBL; M31548; AAA36797.1; JOINED.
 DR EMBL; M31549; AAA36797.1; JOINED.
 DR EMBL; M31550; AAA36797.1; JOINED.
 DR EMBL; BC012609; AAH12609.1; -;
 DR EMBL; A21238; CAA01535.1; -;
 DR EMBL; A21254; CAA01539.1; -;
 DR FIR; A32853; A32853.
 DR FDB; IBY7; 24-OCT-99.
 DR PDB; 1JRR; 18-DEC-02.
 DR Aarhus/Ghent-2DPAGE; 6314; IEP.
 DR Genew; HGNC:8584; SERPINB2.
 DR MIM; 173390; -;
 DR GO; GO:0006916; P:anti-apoptosis; TAS.
 DR InterPro; IPR00215; Serpin.
 DR Pfam; PF00079; serpin; 1.
 DR SMART; SM00093; SERPIN; 1.
 DR PROSITE; PS00284; SERPIN; 1.
 KW Serpin; Serine protease inhibitor; Plasma; Plasminogen activation;
 KW Glycoprotein; Signal; 3D-structure; Polymorphism.
 FT SIGNAL 1 ? NOT CLEAVED.
 FT CARBOHYD 75 75 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 115 115 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT ACT_SITE 380 381 REACTIVE BOND.

FT VARIANT 120 120 N -> D (IN dBSNP:6098).
 FT FT/FTIG=VAR_011743.
 FT VARIANT 229 229 R -> H (IN dBSNP:6100).
 FT FT/FTIG=VAR_014173.
 FT VARIANT 404 404 N -> K (IN dBSNP:6103).
 FT FT/FTIG=VAR_011744.
 FT VARIANT 413 413 S -> C (IN dBSNP:6104).
 FT FT/FTIG=VAR_011745.
 FT CONFLICT 170 170 N -> Y (IN REF. 7).
 FT HELIX 4 22
 FT STRAND 28 30
 FT HELIX 32 44
 FT TURN 45 45
 FT HELIX 48 57
 FT TURN 58 63
 FT HELIX 100 102
 FT HELIX 103 114
 FT STRAND 123 132
 FT TURN 133 134
 FT STRAND 137 137
 FT HELIX 139 149
 FT STRAND 154 156
 FT HELIX 158 176
 FT TURN 177 179
 FT TURN 182 183
 FT TURN 187 188
 FT TURN 192 193
 FT STRAND 196 204

Query Match 3.4%; Score 13; DB 1; Length 415;
 Best Local Similarity 100.0%; Pred. No. 2.6e-05;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 MVLNNAVYFKGW 181
 |||||
 DB 196 MVLNNAVYFKGW 208

RESULT 4

ID PA12_MOUSE STANDARD; PRT; 415 AA.
 AC P12389; Q35687; Q9QWP6; Q9QWP7; Q9QWP8; Q9QWP9; Q9QWQ; Q9QWZ5;
 AC Q9QWZ6;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Plasminogen activator inhibitor-2, macrophage (PAI-2).
 GN SERPINE2 OR PAI2 OR PLAH2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE OF 1-27 FROM N.A.
 RC STRAIN=AKR; TISSUE=Myeloid;
 RX MEDLINE=90059920; PubMed=2583099;
 RA Belin D., Wohlwend A., Schleuning W.-D., Kruthof E.K.O.,
 RA Vasalli J.-D.,
 RT "Facultative polypeptide translocation allows a single mRNA to encode
 the secreted and cytosolic forms of plasminogen activators inhibitor
 2.";
 RT 2.;
 RN EMBL J. 8:3287-3294 (1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AKR; TISSUE=Myeloid;
 RA Belin D.;
 RN Submitted (SEP-1989) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A., AND VARIANTS GUJ-239; TYR-258; ILB-259 AND
 RP AEG-284.
 RC STRAIN=129; TISSUE=Liver;
 RA Belin D., Tapparel C., Sappino N., Silva F.;
 RT "Sequence and tissue-specific expression of the murine PAI-2 gene.";

RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 [4]
 RN SEQUENCE OF 1-34; 54-74; 89-115; 124-156; 163-237 AND 265-302 FROM
 RP N.A.
 RC STRAIN=129/Sv;
 RX MEDLINE=99110951; PubMed=9892694;
 RA Dougherty K.M., Pearson J.M., Yang A.Y., Westrick R.J., Baker M.S.,
 RA Ginsburg D.;
 RT "The plasminogen activator inhibitor-2 gene is not required for normal
 murine development or survival.";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:686-691 (1999).
 CC !- FUNCTION: PAI-2 inhibits urokinase-type plasminogen activator. The
 CC monocyte derived PAI-2 is distinct from the endothelial cell-
 CC derived PAI-1. Not required for normal murine development or
 CC survival.
 CC !- SUBCELLULAR LOCATION: CYTOPLASMIC OR EXTRACELLULAR.
 CC !- PTM: THE SIGNAL SEQUENCE IS NOT CLEAVED.
 CC !- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL; X16490; CAA34507.1; -;
 DR EMBL; AJ000386; CAA04047.1; -;
 DR EMBL; AJ000384; CAA04047.1; JOINED.
 DR EMBL; AJ000385; CAA04047.1; JOINED.
 DR EMBL; AF069683; AAD08915.1; -;
 DR EMBL; AF069685; AAD08916.1; -;
 DR EMBL; AF069684; AAD08916.1; JOINED.
 DR EMBL; AF069687; AAD08917.1; -;
 DR EMBL; AF069686; AAD08917.1; JOINED.
 DR EMBL; AF069689; AAD08918.1; -;
 DR EMBL; AF069688; AAD08918.1; JOINED.
 DR EMBL; AF069690; AAD08919.1; -;
 DR EMBL; AF069692; AAD08920.1; -;
 DR EMBL; AF069691; AAD08920.1; JOINED.
 DR EMBL; AF069694; AAD08921.1; -;
 DR EMBL; AF069693; AAD08921.1; JOINED.
 DR HSP; F05120; I8Y7.
 DR MGD; MGI:97609; Serpinb2.
 DR InterPro; IPR000240; Maspin.
 DR InterPro; IPR000215; Serpin.
 DR Pfam; PF00079; serpin_1.
 DR PRINTS; PR00676; MASPIN.
 DR SMART; SM00093; SERPIN; 1.
 DR PROSITE; PS00284; SERPIN; 1.
 KW Serpin; Serine protease inhibitor; Plasma; Plasminogen activation;
 KW Glycoprotein; Signal; Polymorphism.
 FT SIGNAL 1; NOT CLEAVED.
 FT CARBOHYD 23 23 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 75 75 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 261 261 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT ACT SITE 380 381 REACTIVE BOND.
 FT VARIANT 239 239 A -> E.
 FT VARIANT 258 258 H -> Y.
 FT VARIANT 259 259 T -> I.
 FT VARIANT 284 284 S -> R.
 FT CONFLICT 55 55 A -> T (IN REF. 3 AND 4).
 SQ SEQUENCE 415 AA; 46291 MW; C4FC78C5C3CA7D8B CRC64;

Query Match 3.4%; Score 13; DB 1; Length 415;

Best Local Similarity 100.0%; Pred. No. 2.6e-05;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 MVLNNAVYFKGW 181
 |||||
 DB 196 MVLNNAVYFKGW 208

```

RESULT 5
GDN_MOUSE
ID GDN_MOUSE STANDARD; PRT; 397 AA.
AC Q07235; Q92117;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE GLIA derived nexin precursor (GDN) (Protease nexin I) (PN-1) (Serine
DE protease-inhibitor-4).
GN SERPINE2 OR PI7 OR PNI OR SPI4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Belin D.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=2238257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stopleton M., Soares M.B., Bontaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.N.,
RA Butterfield Y.S.N., Kryzyski M.L., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP SEQUENCE OF 1-366 FROM N.A.
RC STRAIN=BA13/c;
RX MEDLINE=93259128; PubMed=8491179;
RA Vassalli J.-D., Huarte J., Bosco D., Sappino A.P., Sappino N.,
RA Velardi A., Wohlwend A., Erno H., Monard D., Belin D.;
RT "Protease-nexin I as an androgen-dependent secretory product of the
RT murine seminal vesicle."
RL EMBO J. 12:1871-1898 (1993).
CC
CC -!- FUNCTION: THIS GLYCOPROTEIN PROMOTES NEURITE EXTENSION AND IS A
CC SERINE PROTEASE INHIBITOR WITH ACTIVITY TOWARD THROMBIN, TRYPSIN,
CC AND UROKINASE. BINDS HEPARIN.
CC
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC
CC -!- TISSUE SPECIFICITY: MOST ABUNDANT IN SEMINAL VESICLES.
CC
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X70236; CAA49777.1; -
CC EXPL; BC010675; AAH10675.1; -
CC EMBL; X70946; CAA50285.1; -
CC PIR; I48717; I48717.
CC HSSP; P05121; 1A7C.
CC MGD; MGI:101780; Serpine2.

DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; serpin; 1.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
KW Serine protease inhibitor; Serpin; Heparin-binding; Neurone;
KW Glycoprotein; Signal.
FT SIGNAL 1 19
FT CHAIN 20 397
FT CARBOHYD 159 159
FT ACT SITE 364 365
FT CONFLICT 313 313
FT SEQUENCE 397 AA; 44207 MW; B8EAE3CA899D45 CRC64;
SQ
Query Match 2.6%; Score 10; DB 1; Length 397;
Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 170 VLVNAVYFKG 179
Db 182 VLVNAVYFKG 191

RESULT 6
GDN_RAT
ID GDN_RAT STANDARD; PRT; 397 AA.
AC P07092;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE GLIA derived nexin precursor (GDN) (Protease nexin I) (PN-1).
GN SERPINE2 OR PI7 OR PNI.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89107544; PubMed=3427015;
RA Sommer J., Gloor S.M., Rovelli G.F., Hofsteenge J., Nick H.,
RA Veier R., Monard D.;
RT "cDNA sequence coding for a rat glia-derived nexin and its homology
RT to members of the serpin superfamily."
RL Biochemistry 26:6407-6410 (1987).
CC
CC -!- FUNCTION: THIS GLYCOPROTEIN PROMOTES NEURITE EXTENSION AND IS A
CC SERINE PROTEASE INHIBITOR WITH ACTIVITY TOWARD THROMBIN, TRYPSIN,
CC AND UROKINASE. BINDS HEPARIN.
CC
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M17784; AAA41209.1; -
CC PIR; B27496; B27496.
CC HSSP; P05121; 1A7C.
CC
CC InterPro; IPR000215; Serpin.
CC Pfam; PF00079; serpin; 1.
CC SMART; SM00093; SERPIN; 1.
CC PROSITE; PS00284; SERPIN; 1.
KW Serine protease inhibitor; Serpin; Heparin-binding; Neurone;
KW Glycoprotein; Signal.
FT SIGNAL 1 19
FT CHAIN 20 397
FT CARBOHYD 159 159
FT ACT SITE 364 365
FT SEQUENCE 397 AA; 44063 MW; 11EF0790C7297646 CRC64;
SQ
Query Match 2.6%; Score 10; DB 1; Length 397;

```

```

Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 170 VLVNAVYFKG 179
Db 182 VLVNAVYFKG 191

RESULT 7
GDN_HUMAN
ID GDN_HUMAN STANDARD; PRT; 398 AA.
AC P07093;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Gla derived nexin precursor (GDN) (Protease nexin I) (FN-1)
DE (Protease inhibitor 7).
GN SERPINE2 OR P17 OR PNI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88107544; PubMed=3427015;
RA Sommer J., Gloor S.M., Rovelli G.F., Hofsteenge J., Nick H.,
RA Meier R., Monard D.;
RT "cDNA sequence coding for a rat glia-derived nexin and its homology
RT to members of the serpin superfamily.";
RL Biochemistry 26:6407-6410(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87051740; PubMed=2877744;
RA Gloor S.M., Odink K., Guenther J., Nick H., Monard D.;
RT "A glia-derived neurite promoting factor with protease inhibitory
RT activity belongs to the protease nexins.";
RL Cell 47:687-693(1986).
RN [3]
RP SEQUENCE FROM N.A.
RX McGrogan M., Kennedy J., Li M.P., Hsu C., Scott R.W., Simonsen C.C.,
RA Baker J.B.;
RT "Molecular cloning and expression of two forms of human protease
RT nexin I.";
RL Biotechnology 6:172-177(1988).
RN [4]
RP SEQUENCE FROM N.A.
RC TIGR=Skin;
RX MEDLINE=22389257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.N., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Muzny K.C., Hale S., Garcia A.M., Gay L.J., Hultyk S.W.,
RA Fahey J., Heaton E., Kettaman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smalius D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP SEQUENCE OF 20-47, AND CHARACTERIZATION.
RX MEDLINE=85207723; PubMed=3997857;
RA Scott R.W., Bergman B.L., Bajpai A., Herish R.T., Rodriguez H.,
RA Jones B.N., Barreda C., Watts S., Baker J.B.;
RT "Protease nexin. Properties and a modified purification procedure.";
RL J. Biol. Chem. 260:7029-7034(1985).
CC -!- FUNCTION: THIS GLYCOPROTEIN PROMOTES NEURITE EXTENSION AND IS A
CC SERINE PROTEASE INHIBITOR WITH ACTIVITY TOWARD THROMBIN, TRYPSIN,
CC AND UROKINASE. BINDS HEPARIN.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- MISCELLANEOUS: TWO VARIANTS ARE DERIVED FROM TWO DIFFERENT
CC SUBCLONES.
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M17783; AAA35883.1; -.
CC EMBL; BC015663; AAH15663.1; -.
CC PIR; A37274; A37274.
CC KSP; P05121; IA706.
CC Genew; HGNC:8951; SERPINE2.
CC MIM; 177010; -.
CC GO; GO:0005576; C:extracellular; NAS.
CC GO; GO:0008201; F:heparin binding activity; NAS.
CC GO; GO:0004867; F:serine protease inhibitor activity; NAS.
CC GO; GO:0003034; P:regulation of cell migration; NAS.
CC GO; GO:0003162; P:regulation of proteolysis and peptidolysis; NAS.
CC InterPro; IPR000215; Serpin.
CC Pfam; PF00079; serpin; 1.
CC SMART; SM00093; SERPIN; 1.
CC PROSITE; PS00284; SERPIN; 1.
CC Serine protease inhibitor; Serpin; Heparin-binding; Neurone;
KW Glycoprotein; Signal.
FT SIGNAL 1 19
FT CHAIN 20 398 GLIA DERIVED NEXIN.
FT CARBOHYD 118 118 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 159 159 N-LINKED (GLCNAC...) (POTENTIAL).
FT ACT_SITE 365 365 REACTIVE BOND (POTENTIAL).
FT VARIANT 329 330 TG -> R.
FT FTID=VAR_007098.
FT CONFLICT 261 261 S -> E (IN REF. 2).
SQ SEQUENCE 398 AA; 44002 MW; 2A165604E2CB658 CRC64;

Query Match 2.6%; Score 10; DB 1; Length 398;
Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 170 VLVNAVYFKG 179
Db 182 VLVNAVYFKG 191

RESULT 8
SI21_RAT
ID SI21_RAT STANDARD; PRT; 214 AA.
AC P09005;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-NOV-1988 (Rel. 09, Last annotation update)
DE Serine protease inhibitor 2.1 (SPI-2.1) (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87144617; PubMed=3493437;
RA Hill R.E., Hastie N.D.;
RT "Accelerated evolution in the reactive centre regions of serine
RT protease inhibitors.";
RL Nature 326:96-99(1987).

```


CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X13149; CAA31547.1; -.
 DR HSSP; P01011; IAS4.
 DR InterPro; IPR000215; Serpin.
 DR Pfam; PF00079; serpin; 1.
 DR SMART; SM00093; SERPIN; 1.
 DR PROSITE; PS00284; SERPIN; 1.
 DR Serpin; Serine protease inhibitor.
 KW NON_TER 1
 FT ACT SITE 180 181 REACTIVE BOND (BY SIMILARITY).
 SQ SEQUENCE 214 AA; 24218 MW; B1F4CE59425BAAF CRC64;
 Query Match 2.4%; Score 9; DB 1; Length 214;
 Best Local Similarity 100.0%; Pred. No. 0.21;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 333 GTEATATG 341
 DB 166 GTEATATG 174
 RESULT 9
 OVAX CHICK
 ID OVAX CHICK STANDARD; PRT; 232 AA.
 AC P01073;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-APR-1993 (Rel. 25, Last annotation update)
 DE Gene X protein (Ovalbumin-related) (Fragment).
 GN X.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6102623; PubMed=7418002;
 RA Heilig R., Ferrin F., Gannon F., Mandel J.L., Chambon P.;
 RT "The ovalbumin gene family: structure of the X gene and evolution of
 RT duplicated split genes.";
 RL Cell 20:625-637(1980).
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; J00920; AAA68881.1; -.
 DR EMBL; J00918; AAA68881.1; JOINED.
 DR EMBL; J00919; AAA68881.1; JOINED.
 DR EMBL; V00385; CAA23683.1; -.
 DR EMBL; V00386; CAA23684.1; -.
 DR EMBL; V00387; CAA23685.1; -.
 DR PIR; A01243; DXCH.
 DR HSSP; P01012; LOVA.
 DR InterPro; IPR000215; Serpin.
 DR Pfam; PF00079; serpin; 1.
 DR SMART; SM00093; SERPIN; 1.
 DR PROSITE; PS00284; SERPIN; 1.

KW Serpin.
 FT NON_TER 1
 SQ SEQUENCE 232 AA; 26291 MW; 6B5B86EC4D3B9195 CRC64;
 Query Match 2.4%; Score 9; DB 1; Length 232;
 Best Local Similarity 100.0%; Pred. No. 0.22;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 355 FRAHPPLF 363
 DB 205 FRAHPPLF 213
 RESULT 10
 PT16 BOVIN
 ID PT16 BOVIN STANDARD; PRT; 378 AA.
 AC O02739;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Serine proteinase inhibitor B-43.
 CS Bos taurus (Bovine).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 CC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=97165993; PubMed=9013786;
 RA Nakaya N., Nishibori M., Kawabata M., Saeki K.;
 RT "Cloning of a serine proteinase inhibitor from bovine brain:
 RT expression in the brain and characterization of its target
 RT proteinases.";
 RL Brain Res. Mol. Brain Res. 42:293-300(1996).
 CC -!- FUNCTION: MAY BE INVOLVED IN THE REGULATION OF SERINE PROTEINASES
 CC PRESENT IN THE BRAIN OR EXTRAVASATED FROM THE BLOOD.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- TISSUE SPECIFICITY: BRAIN.
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D55670; BAA19875.1; -.
 DR HSSP; P05120; IBY7.
 DR InterPro; IPR000215; Serpin.
 DR Pfam; PF00079; serpin; 1.
 DR SMART; SM00093; SERPIN; 1.
 DR PROSITE; PS00284; SERPIN; 1.
 KW Serpin; Serine protease inhibitor.
 FT ACT SITE 343 344 REACTIVE BOND (BY SIMILARITY).
 SQ SEQUENCE 378 AA; 42560 MW; 664F499CCFCE263A CRC64;
 Query Match 2.4%; Score 9; DB 1; Length 378;
 Best Local Similarity 100.0%; Pred. No. 0.34;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 328 EVTEGTEA 336
 DB 324 EVTEGTEA 332
 RESULT 11
 OVAL COTUA
 ID OVAL COTUA STANDARD; PRT; 382 AA.
 AC P19164;
 DT 01-NOV-1990 (Rel. 16, Created)

DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 01-APR-1993 (Rel. 25, Last annotation update)
 DE Ovalbumin.
 OS Coturnix coturnix japonica (Japanese quail).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Coturnix.
 OX NCBI_TaxID=93934;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Oviduct;
 RX MEDLINE=9101650; PubMed=2216734;
 RA Mucha J., Klaudityova V., Hames J., Simuth J.;
 RT "The sequence of Japanese quail ovalbumin cDNA."
 RL Nucleic Acids Res. 18:5553-5553(1990).
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X53964; CAA37916.1; --
 DR PIR; S11433; S11433.
 DR HSSP; P01012; IOVA.
 DR InterPro; IPR000215; Serpin.
 DR Pfam; PF000079; serpin; 1.
 DR SMART; SM00093; SERPIN; 1.
 DR PROSITE; PS00284; SERPIN; 1.
 KW Serpin; Glycoprotein.
 FT INIT MET 0 0
 FT DISULFID 73 120 BY SIMILARITY.
 FT CARBOHYD 292 292 N-LINKED (GLNAC. . .) (POTENTIAL).
 FT CARBOHYD 311 311 N-LINKED (GLNAC. . .) (POTENTIAL).
 FT ACT SITE 352 353 REACTIVE BOND HOMOLOG.
 SQ SEQUENCE 382 AA; 42108 MW; A4BB59A1BA5EF316 CRC64;

 Query Match 2.4%; Score 9; DB 1; Length 382;
 Best Local Similarity 100.0%; Pred. No. 0.35;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 355 FRADHPFLF 363
 |||||
 DB 355 FRADHPFLF 363

 RESULT 12.
 OVAL_CHICK STANDARD; PRT; 385 AA.
 AC P01012;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=42069038; PubMed=6272839;
 RA Woo S.I.C., Beattie W.G., Catterall J.F., Dugaiczky A., Staden R.,
 RA Brownlee G.G., O'Malley B.W.;
 RT "Complete nucleotide sequence of the chicken chromosomal ovalbumin
 RT gene and its biological significance."
 RL Biochemistry 20:6437-6446(1981).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=78199842; PubMed=661981;

RA McReynolds L., O'Malley B.W., Nisbet A.D., Fothergill J.E., Givol D.,
 RA Fields S., Robertson M., Brownlee G.G.;
 RT "Sequence of chicken ovalbumin mRNA."
 RL Nature 273:723-728(1978).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=79010682; PubMed=692731;
 RA Catterall J.F., O'Malley B.W., Robertson M.A., Staden R.,
 RA Tanaka Y., Brownlee G.G.;
 RT "Nucleotide sequence homology at 12 intron-exon junctions in the
 RT chick ovalbumin gene."
 RL Nature 275:510-513(1978).
 RN [4]
 RP SEQUENCE OF 1-35.
 RX MEDLINE=78116057; PubMed=272676;
 RA Palmier R.D., Gagnon J., Walsh K.A.;
 RT "Ovalbumin: a secreted protein without a transient hydrophobic leader
 RT sequence."
 RL Proc. Natl. Acad. Sci. U.S.A. 75:94-98(1978).
 RN [5]
 RP SEQUENCE OF 1-16.
 RX MEDLINE=79186958; PubMed=751625;
 RA Thompson E.O.P., Fisher W.K.;
 RT "A correction and extension of the acetylated amino terminal sequence
 RT of ovalbumin."
 RL Aust. J. Biol. Sci. 31:443-446(1978).
 RN [6]
 RP SEQUENCE OF 5-16; 29-35; 60-78; 115-123; 366-373 AND 379-385.
 RX MEDLINE=79186957; PubMed=751624;
 RA Thompson E.O.P., Fisher W.K.;
 RT "Amino acid sequences containing half-cystine residues in ovalbumin."
 RL Aust. J. Biol. Sci. 31:433-442(1978).
 RN [7]
 RP N-TERMINUS ACETYLATION, AND PHOSPHORYLATION OF SER-68; SER-236 AND
 RP SER-240.
 RX MEDLINE=22056091; PubMed=12060738;
 RA MacCoss M.J., McDonald W.H., Saraf A., Sadygov R., Clark J.M.,
 RA Tascio J.J., Gould K.L., Wolters D., Washburn M., Weiss A., Clark J.I.,
 RA Yates J.R., III;
 RT "Shotgun identification of protein modifications from protein
 RT complexes and lens tissue."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:7900-7905(2002).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (1.95 ANGSTROMS).
 RX MEDLINE=90370102; PubMed=2395463;
 RA Stein P.E., Leslie A.G.W., Finch J.T., Turnell W.G., McLaughlin P.J.,
 RA Carrell R.W.;
 RT "Crystal structure of ovalbumin as a model for the reactive centre of
 RT serpins."
 RL Nature 347:99-102(1990).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (1.95 ANGSTROMS).
 RX MEDLINE=92046044; PubMed=1942038;
 RA Stein P.E., Leslie A.G.W., Finch J.T., Carrell R.W.;
 RT "Crystal structure of uncleaved ovalbumin at 1.95-A resolution."
 RL J. Mol. Biol. 221:941-959(1991).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
 RX MEDLINE=90278960; PubMed=2352279;
 RA Wright H.T., Qian H.X., Huber R.;
 RT "Crystal structure of plakalbumin, a proteolytically nicked form of
 RT ovalbumin. Its relationship to the structure of cleaved alpha-1-
 RT proteinase inhibitor."
 RL J. Mol. Biol. 213:513-528(1990).
 RN [11]
 RP REVIEW.
 RX MEDLINE=21312433; PubMed=11419711;
 RA Huntington J.A., Stein P.E.;
 RT "Structure and properties of ovalbumin."
 RL J. Chromatogr. B 756:189-198(2001).
 CC -!- FUNCTION: Not known.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Major protein of egg white.

CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 CC -!- DATABASE: NAMS=Washington enzyme manual;
 CC WWW="http://www.worthington-biochem.com/manual/OA.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; J00895; AAB59956.1; -;
 CC EMBL; V00438; CAA23716.1; -;
 CC EMBL; V00383; CAA23682.1; -;
 CC EMBL; M34352; AAA48998.1; -;
 CC EMBL; M34346; AAA48998.1; JOINED.
 CC EMBL; M34347; AAA48998.1; JOINED.
 CC EMBL; M34348; AAA48998.1; JOINED.
 CC EMBL; M34349; AAA48998.1; JOINED.
 CC EMBL; M34350; AAA48998.1; JOINED.
 CC EMBL; M34351; AAA48998.1; JOINED.
 CC PIR; A90455; OACH.
 CC PDB; IOVA; 1S-JUL-92.
 CC PDB; IUT1; 1S-DEC-02.
 CC PDB; 1VAC; 20-JUN-96.
 CC DR GlycoSuiteDB, P01012, -;
 CC DR InterPro; IPR000215; Serpin.
 CC DR Pfam; PF000079; serpin; 1.
 CC DR SMART; SMO00093; SERPIN; 1.
 CC DR PROSITE; PS00284; SERPIN; 1.
 CC KW Serpin; Acetylation; Phosphorylation; Glycoprotein; 3D-structure;
 CC Allersen.
 CC XT INIT MET 0 0
 CC FT MOD RES 1 1
 CC FT MOD RES 68 68
 CC FT DISULFID 73 120
 CC FT MOD RES 236 236
 CC FT MOD RES 240 240
 CC FT CARBOHYD 292 292
 CC FT MOD RES 344 344
 CC FT ACT SITE 352 353
 CC FT VARIANT 311 311
 CC FT CONFLICT 187 187
 CC FT HELIX 3 21
 CC FT TURN 23 24
 CC FT STRAND 27 29
 CC FT HELIX 31 43
 CC FT TURN 44 44
 CC FT HELIX 47 57
 CC FT TURN 58 58
 CC FT TURN 60 61
 CC FT TURN 63 64
 CC FT HELIX 67 70
 CC FT TURN 71 75
 CC FT TURN 79 80
 CC FT HELIX 81 91
 CC FT STRAND 97 108
 CC FT TURN 109 110
 CC FT STRAND 113 113
 CC FT HELIX 115 124
 CC FT STRAND 129 132
 CC FT TURN 135 137
 CC FT HELIX 138 152
 CC FT TURN 153 155
 CC FT TURN 163 164
 CC FT TURN 168 169
 CC FT STRAND 172 182
 CC FT STRAND 184 186
 CC FT HELIX 190 192
 CC FT STRAND 194 199
 CC FT STRAND 205 222
 CC FT TURN 223 226

FT STRAND 227 234
 FT TURN 235 236
 FT STRAND 239 246
 FT TURN 249 251
 FT HELIX 252 258
 FT TURN 261 267
 FT TURN 268 268
 FT STRAND 270 272
 FT STRAND 274 283
 FT STRAND 285 292
 FT HELIX 293 299
 FT TURN 300 301
 FT HELIX 304 306
 FT TURN 308 309
 FT TURN 313 315
 FT STRAND 316 316
 FT TURN 319 320
 FT STRAND 325 334
 FT STRAND 338 340
 FT STRAND 357 359
 FT STRAND 364 370
 FT TURN 371 373
 FT STRAND 376 382
 SQ SEQUENCE 385 AA; 42750 MW; 52B4339642388FE3 CRC64;
 Query Match 2.4%; Score 9; DB 1; Length 385;
 Best Local Similarity 100.0%; Pred. No. 0.35;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 355 FRADHPFLP 363
 D5 358 FRADHPFLP 366
 |||||
 RESULT 13
 ID OVAY CHICK STANDARD; PRT; 388 AA.
 AC P01014;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-APR-1993 (Rel. 25, Last annotation update)
 DE Gene Y protein (Ovalbumin-related).
 GN Y.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83014329; PubMed=7122240;
 RA Heilig R., Muraskowsky R., Kloeffer C., Mandel J.L.;
 RT "The ovalbumin gene family: complete sequence and structure of the Y
 RT gene."
 RL Nucleic Acids Res. 10:4363-4382(1982).
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; J00922; AAA68882.1; -;
 CC PIR; A01244; DYCH.
 CC HSP; P01012; ICVA.
 CC InterPro; IPR000215; Serpin.
 CC Pfam; PF00079; serpin; 1.
 CC SMART; SMO0093; SERPIN; 1.
 CC PROSITE; PS00284; SERPIN; 1.
 CC KW Serpin.

FT ACT_SITE 353 354 REACTIVE BOND HOMOLOG.
FT CARBOHYD 95 95 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 215 215 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 312 312 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 388 AA; 43772 MW; 2AF17B715D7F461 CRC64;

Query Match 2.4%; Score 9; DB 1; Length 388;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 355 FRADHPF.F 363
Db 361 FRADHPF.F 369

RESULT 14
SB12 HUMAN
ID SB12 HUMAN STANDARD; PRT; 405 AA.
AC Q96P63;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serpin B2.
GN SERPINB2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=21623618; PubMed=11604408;
RA Askew Y.S., Pak S.C., Luke C.J., Askew D.J., Cataltepe S., Mills D.R.,
Kato H., Lehoczy J., Dewar K., Birren B., Silverman G.A.;
RT "SERPINB2 is a novel member of the human ov-serpin family that is
widely expressed and inhibits trypsin-like serine proteinases.";
RL J. Biol. Chem. 276:49320-49330(2001).
CC -!- FUNCTION: Inhibits trypsin and plasmin, but not thrombin,
coagulation factor Xa, or urokinase-type plasminogen activator.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC -!- TISSUE SPECIFICITY: Expressed in many tissues, including brain,
bone marrow, lymph node, heart, lung, liver, pancreas, testis,
ovary, and intestine.
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; AF411191; AAL05571.1; -;
CC Genew; HGNC:14220; SERPINB2.
CC GO; GO:0005737; Cytoplasm; NAS.
CC GO; GO:0019899; F:enzyme binding activity; NAS.
CC GO; GO:0004868; F:serpin; IMP.
CC GO; GO:0030304; F:trypsin inhibitor activity; IMP.
CC GO; GO:0042177; P:negative regulation of protein catabolism; NAS.
CC InterPro; IPR000215; Serpin.
CC Pfam; PF00079; Serpin; 1.
CC SMART; SMC0093; SERPIN; 1.
CC PROSITE; PS00284; SERPIN; 1.
KW Serpin; Serine protease inhibitor.
FT ACT SITE 370 371 REACTIVE BOND (BY SIMILARITY).
SQ SEQUENCE 405 AA; 46276 MW; FFEL2D4C9B7F3DFA CRC64;

Query Match 2.4%; Score 9; DB 1; Length 405;
Best Local Similarity 100.0%; Pred. No. 0.37;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 VLNAVYFK 178

Db 186 VLNAVYFK 194

RESULT 15
CPI3_RAT
ID CPI3_RAT STANDARD; PRT; 413 AA.
AC P05544;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Contrapsin-like protease inhibitor 3 precursor (CPI-23) (Serine
protease inhibitor 1) (SPI-1).
DE Rattus norvegicus (Rat).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Liver;
RX MEDLINE=91324305; PubMed=1864837;
RA Ohkubo K., Ogata S., Misumi Y., Takami N., Ikebara Y.;
RT "Molecular cloning and characterization of rat contrapsin-like
protease inhibitor and related proteins.";
RL J. Biochem. 109:243-250(1991).
RN [2]
RP SEQUENCE OF 11-413 FROM N.A.
RX TISSUE=Liver;
RX MEDLINE=90306038; PubMed=1694763;
RA Pages G., Rouayenc J.F., le Cam G., Mariller M., le Cam A.;
RT "Molecular characterization of three rat liver serine-protease
inhibitors affected by inflammation and hypophysectomy. Protein and
mRNA analysis and cDNA cloning.";
RL Eur. J. Biochem. 190:385-391(1990).
RN [3]
RP SEQUENCE OF 82-403 FROM N.A.
RX TISSUE=Liver;
RX MEDLINE=87166046; PubMed=3494016;
RA Yoon J.-B., Towle H.C., Seelig S.;
RT "Growth hormone induces two mRNA species of the serine protease
inhibitor gene family in rat liver.";
RL J. Biol. Chem. 262:4284-4289(1987).
CC -!- FUNCTION: INHIBITS TRYPSIN, BUT NOT CHYMOTRYPSIN OR
ELASTASE.
CC -!- TISSUE SPECIFICITY: LIVER.
CC -!- INDUCTION: BY GROWTH HORMONE.
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
CC -!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-11 IS THE INITIATOR.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; D00752; BAA00649.1; -;
CC EMBL; X16357; CAA34406.1; -;
CC EMBL; M15917; AAA42172.1; -;
CC FIR; S08102; S08102.
CC HSP; P01011; ZACH.
CC InterPro; IPR000215; Serpin.
CC Pfam; PF00079; serpin; 1.
CC SMART; SMC0093; SERPIN; 1.
CC PROSITE; PS00284; SERPIN; 1.
KW Serpin; serine protease inhibitor; Glycoprotein; Signal.
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 413 CONTRAPSIN-LIKE PROTEASE INHIBITOR 3.
FT ACT SITE 379 380 REACTIVE BOND (BY SIMILARITY).
FT CARBOHYD 102 102 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 182 182 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 220 220 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 267 267 N-LINKED (GLNAC...) (POTENTIAL);
 FT CONFLICT 245 245 V -> L (IN REF. 3).
 SQ SEQUENCE 413 AA; 46277 MW; 3AAACFAFAA9BE7 CRC64;
 Query Match 2.4%; Score 9; DB 1; Length 413;
 Best Local Similarity 100.0%; Pred. No. 0.37;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 333 GTEATATG 341
 Db 365 GTEATATG 373
 |||||

Search completed: December 12, 2003, 16:33:54
 Job time : 24 secs

GenCore version 5.1.6
Copyright (C) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 16:31:43 ; Search time 21 Seconds
(without alignments)
1740.195 Million cell updates/sec

Title: US-09-508-997A-2
Perfect score: 380
Sequence: 1 MASLAAANAEFCNLFREMD.....FLFVIRKDDIIILFSKGVSCP 350

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR_76:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	13	3.4	415	2 S20047	plasminogen activa
2	13	3.4	415	2 A32853	plasminogen activa
3	10	2.6	238	2 T30945	hypothetical prote
4	10	2.6	397	2 I48717	proteinase inhibit
5	10	2.6	397	2 B27496	proteinase inhibit
6	10	2.6	398	2 A37274	glia-derived nexin
7	9	2.4	213	2 A26423	serine proteinase
8	9	2.4	232	1 DXCH	ovalbumin-related
9	9	2.4	248	2 AC1919	hypothetical prote
10	9	2.4	383	2 S11433	ovalbumin - Japane
11	9	2.4	386	1 OACH	ovalbumin [validat
12	9	2.4	388	1 DYCH	ovalbumin-related
13	9	2.4	403	2 S08102	serine proteinase
14	8	2.1	121	2 T31297	hypothetical prote
15	8	2.1	262	2 B83277	hypothetical prote
16	8	2.1	291	2 A30558	B-cell surface ant
17	8	2.1	341	2 H97106	probable permease
18	8	2.1	376	1 A48681	placental thrombin
19	8	2.1	384	2 E97168	probable CDP-4-ret
20	8	2.1	391	2 JC7118	headpin serine pro
21	8	2.1	433	1 ITHUC	alpha-1-antichymot
22	8	2.1	612	2 T35430	probable long-chain
23	8	2.1	613	2 T16885	hypothetical prote
24	8	2.1	626	2 C87672	peptide transporte
25	8	2.1	922	2 A30816	band 3 anion trans
26	8	2.1	1179	2 F71190	probable chromosom
27	7	1.8	19	2 S43652	serpin WSZC1 - whe
28	7	1.8	82	2 AB2351	hypothetical prote
29	7	1.8	121	2 B84770	probable serpin [i

30	7	1.8	123	2 T00549	60S ribosomal prot
31	7	1.8	152	2 AE2333	hypothetical prote
32	7	1.8	157	2 A61017	phosphopyruvate hy
33	7	1.8	157	2 C71060	hypothetical prote
34	7	1.8	167	2 AP2720	hypothetical prote
35	7	1.8	170	2 AE0630	4-hydroxyphenylace
36	7	1.8	188	2 T10859	auxin-induced prot
37	7	1.8	198	2 AC1364	protein gp51 [Bact
38	7	1.8	220	2 D83495	conserved hypothet
39	7	1.8	240	2 S73285	hypothetical prote
40	7	1.8	242	2 C71006	hypothetical prote
41	7	1.8	242	2 T32122	hypothetical prote
42	7	1.8	254	2 C82868	plasmid replicatio
43	7	1.8	261	2 T20264	hypothetical prote
44	7	1.8	266	2 C75613	probable molybdenu
45	7	1.8	270	2 B69781	multidrug-efflux t

ALIGNMENTS

RESULT 1

S20047

Plasminogen activator inhibitor 2 - mouse

C:Species: Mus musculus (house mouse)

C>Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 16-Jul-1999

C:Accession: S20047; S09616

R:Belin, D.

submitted to the EMBL Data Library, September 1989

A:Reference number: S20047

A:Accession: S20047

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-415 <BEL>

A:Cross-references: EMBL:X16490; NID:953589; PIDN:CAA34507.1; PID:G53590

R:Belin, D.; Wohlwend, A.; Schleuning, W.D.; Kruthof, E.K.O.; Vassalli, J.D.

EMBO J. 8, 3287-3294, 1989

A:Title: Facultative polypeptide translocation allows a single mRNA to encode the se

A:Reference number: S09615; MUID:90059920; PMID:2583099

A:Accession: S09616

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-27 <BEL2>

A:Cross-references: EMBL:X16490

C:Superfamily: antithrombin III

Query Match 3.4%; Score 13; DB 2; Length 415;

Best Local Similarity 100.0%; Pred. No. 6.2e-05;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 169 MVLNAVYFKGW 181

Db 136 MVLNAVYFKGW 208

RESULT 2

A32853

plasminogen activator inhibitor 2 precursor - human

N:Alternate names: urokinase inhibitor

C:Species: Homo sapiens (man)

C>Date: 22-Nov-1999 #sequence_revision 22-Nov-1999 #text_change 20-Oct-2000

C:Accession: A32853; I34218; A31366; A46543; A29362; A26553; A53815; A25021; S02435;

J. Biol. Chem. 264, 5495-5502, 1989

A:Title: Structure of the gene for human plasminogen activator inhibitor-2. The near

A:Reference number: A32853; MUID:89174589; PMID:2494165

A:Accession: A32853

A:Molecule type: DNA

A:Residues: 1-415 <YER>

A:Cross-references: GB:M24651; GB:M24652; GB:M24653; GB:M24654; GB:M24655; GB:M24656

R:Samia, J.A.; Alexander, S.J.; Horton, K.W.; Byers, M.G.; Shows, T.B.

Genomics 6, 159-167, 1990

A:Title: Chromosomal organization and localization of the human urokinase inhibitor

A:Reference number: I54218; MUID:90152578; PMID:2303256
 A:Accession: I54218
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-415 <RES>
 A:Cross-references: GB:M31551; NID:G340152; PIDN:AAA6797.1; PID:G340154
 R:Antalis, T.M.; Clark, M.A.; Barnes, T.; Lehrbach, P.R.; Devine, P.L.; Schvezov, G.; G.
 Proc. Natl. Acad. Sci. U.S.A. 85, 985-989, 1988
 A:Title: Cloning and expression of a cDNA coding for a human monocyte-derived plasminogen
 A:Reference number: A31366; MUID:88125032; PMID:3257578
 A:Accession: A31366
 A:Molecule type: mRNA
 A:Residues: 1-415 <ANT>
 A:Cross-references: GB:J03603; NID:G189546; PIDN:AAA60004.1; PID:G189547
 R:Webb, A.C.; Collins, K.L.; Snyder, S.E.; Alexander, S.J.; Roserwasser, L.J.; Eddy, R.I.
 J. Exp. Med. 166, 77-94, 1997
 A:Title: Human monocyte Arg-serpin cDNA. Sequence, chromosomal assignment, and homology
 A:Reference number: A46543; MUID:8725228; PMID:3496414
 A:Accession: A46543
 A:Molecule type: mRNA
 A:Residues: 1-415 <WEB>
 A:Cross-references: GB:Y00630; NID:G35267; PIDN:CAA68666.1; PID:G35268
 R:Schleuning, W.D.; Medcalf, R.L.; Hession, C.; Rothenbuehler, R.; Shaw, A.; Knutthof, E.
 Mol. Cell. Biol. 7, 4564-4567, 1987
 A:Title: Plasminogen activator inhibitor 2: regulation of gene transcription during phor
 A:Reference number: A29362; MUID:88142852; PMID:3325828
 A:Accession: A29362
 A:Molecule type: mRNA
 A:Residues: 1-119, D', 121-403, 'K', 405-412, 'C', 414-415 <SCH>
 A:Cross-references: GB:M18082; NID:G189562; PIDN:AAA60006.1; PID:G189563
 R:Ye, R.D.; Wu, T.C.; Sadler, J.B.
 J. Biol. Chem. 262, 3718-3725, 1987
 A:Title: cDNA cloning and expression in Escherichia coli of a plasminogen activator inh
 A:Reference number: A26553; MUID:87137674; PMID:3029122
 A:Accession: A26553
 A:Molecule type: mRNA
 A:Residues: 1-119, D', 121-403, 'K', 405-412, 'C', 414-415 <YE2>
 A:Cross-references: GB:J02685; NID:G189544; PIDN:AAA36413.1; PID:G189545
 R:Experimental source: placenta
 R:Jensen, P.H.; Schueler, E.; Woodrow, G.; Richardson, M.; Goss, N.; Hojrup, P.; Peterse
 J. Biol. Chem. 269, 15394-15398, 1994
 A:Title: A unique interhelical insertion in plasminogen activator inhibitor-2 contains
 A:Reference number: A53815; MUID:94253109; PMID:7910824
 A:Accession: A53815
 A:Molecule type: protein
 A:Residues: 83-57 <JEN>
 R:Kruithof, E.K.O.; Vassalli, J.D.; Schleuning, W.D.; Mattaliano, R.J.; Bachmann, F.
 J. Biol. Chem. 261, 11207-11213, 1986
 A:Title: Purification and characterization of a plasminogen activator inhibitor from the
 A:Reference number: A25021; MUID:86278222; PMID:3090045
 A:Accession: A25021
 A:Molecule type: protein
 A:Residues: 347-376 <KRU>
 R:Kiso, U.; Kaudewitz, H.; Henschen, A.; Astedt, B.; Kruithof, E.K.O.; Bachmann, F.
 FEBS Lett. 230, 51-56, 1988
 A:Title: Determination of intermediates, products and cleavage site in the reaction betw
 A:Reference number: S02435; MUID:88167197; PMID:3280346
 A:Accession: S02435
 A:Molecule type: protein
 A:Residues: 377-415 <KIS>
 R:Kruithof, E.K.; Cousin, E.
 Biochem. Biophys. Res. Commun. 156, 383-388, 1988
 A:Title: Plasminogen activator inhibitor 2. Isolation and characterization of the promot
 A:Reference number: I52229; MUID:89025873; PMID:2845977
 A:Accession: I52229
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-56 <RE2>
 A:Cross-references: GB:M23092; NID:G189559; PIDN:AAA60005.1; PID:G189561
 C:Genetics:
 A:Gene: GDB:PA12; PLANH2
 A:Cross-references: GDB:120298; OMIM:173390
 A:Map position: 18q21.2-18q22

A:Introns: 56/3; 96/3; 139/3; 179/1; 226/3; 281/3
 C:Superfamily: antithrombin III
 C:Keywords: acetylated amino end; glycoprotein; serine proteinase inhibitor
 P:1-22/Domain: signal sequence #status predicted <SIG>
 P:23-415/Product: plasminogen activator inhibitor 2 #status predicted <YAT>
 P:23/Modified site: acetylated amino end (ser) (in mature form) #status predicted
 P:75,115,339/Binding site: carboxylate (Asn) (covalent) #status predicted
 P:83,84,86/Cross-link: isopeptide (Gln) (interchain to lys N6-amino of unidentified p
 P:380-381/Cleavage site: Arg-Thr (plasminogen activator) #status experimental
 P:380/Inhibitory site: Arg (plasminogen activator) #status experimental
 Query Match 3.4%; Score 13; DB 2; Length 415;
 Best Local Similarity 100.0%; Pred. No. 6.2e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 169 MVLNVAVYFKGK 181
 |||||
 Db 196 MVLNVAVYFKGK 208
 |||||
 RESULT 3
 T30945
 hypothetical protein AC7.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jun-2000
 C:Accession: T30945
 R:Nelson, J.; Hawkins, J.
 submitted to the EMBL Data Library, June 1999
 A:Description: The sequence of C. elegans cosmid AC7.
 A:Reference number: Z20944
 A:Accession: T30945
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-238 <NEL>
 A:Cross-references: EMBL:U61957; PIDN:AA03419.1
 A:Experimental source: strain Bristol N2; clone AC7
 C:Genetics:
 A:Map position: IV
 A:Introns: 10/2; 30/3; 72/3; 159/3
 A:Note: AC7.3
 C:Superfamily: Caenorhabditis elegans hypothetical protein AC7.3
 Query Match 2.6%; Score 10; DB 2; Length 238;
 Best Local Similarity 100.0%; Pred. No. 0.047;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 146 VENETHGKIK 155
 |||||
 Db 92 VENETHGKIK 101
 |||||
 RESULT 4
 I48717
 proteinase inhibitor nexin I precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 16-Jul-1999
 C:Accession: I48717; S70772; S35731
 R:Vassalli, J.D.; Huarte, J.; Bosco, D.; Sappino, A.P.; Velardi, A.; Wol
 EMBL J. 12, 1871-1878, 1993
 A:Title: Protease-nexin I as an androgen-dependent secretory product of the murine se
 A:Reference number: I48717; MUID:93259128; PMID:8491179
 A:Accession: I48717
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: mRNA
 A:Residues: 1-397 <RES>
 A:Cross-references: EMBL:X70296; NID:G551064; PIDN:CAA43777.1; PID:G551065
 A:Accession: S70772
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-86 <VAS>
 A:Cross-references: EMBL:X70946; NID:G57930; PIDN:CAA50285.1; PID:G57931
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1993;
 C:Genetics:

A:Gene: FN-1
C:Superfamily: antithrombin III
C:Keywords: serine proteinase inhibitor
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-397/Product: proteinase inhibitor nexin I #status predicted <NAT>

Query Match 2.6%; Score 10; DB 2; Length 397;
Best Local Similarity 100.0%; Pred. No. 0.076;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 VLVNAVYFKG 179
|||||
Db 182 VLVNAVYFKG 191

RESULT 5
B27496
proteinase inhibitor nexin I precursor - rat (fragment)
N:Alternate names: glia-derived nexin (GDN)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Jun-1988 #sequence revision 30-Jun-1988 #text_change 16-Jul-1999
C:Accession: B27496; A34538; A42351; B42351; C42351
R:Sommer, J.; Gloor, S.M.; Rovelli, G.F.; Hofsteenge, J.; Nick, H.; Meier, R.; Monard, D.
Biochemistry 26, 6407-6410, 1987
A:Title: cDNA sequence coding for a rat glia-derived nexin and its homology to members
A:Reference number: A90519; MUID:88107544; PMID:3427015
A:Accession: B27496
A:Molecule type: mRNA
A:Residues: 1-397 <SOM>
A:Cross-references: GB:M17784; NID:G204283; PIDN:AAA1209.1; PID:G204284
A>Note: the authors translated the codon TGG for residue 156 as Thr
R:Nick, H.; Hofsteenge, J.; Shaw, E.; Rovelli, G.; Monard, D.
Biochemistry 29, 2417-2421, 1990
A:Title: Functional sites of glia-derived nexin (GDN): importance of the site reacting w
A:Reference number: A34538; MUID:90248459; PMID:2337608
A:Accession: A34538
A>Status: preliminary
A:Molecule type: protein
A:Residues: 364-394 <NIC>
R:Rovelli, G.; Stone, S.R.; Guidolin, A.; Sommer, J.; Monard, D.
Biochemistry 31, 3542-3549, 1992
A:Title: Characterization of the heparin-binding site of glia-derived nexin/protease nex
A:Reference number: A42351; MUID:92207980; PMID:1554734
A:Accession: A42351
A:Molecule type: protein
A:Residues: 82-96 <ROV1>
A>Note: sequence extracted from NCBI backbone (NCBIP:93846)
A:Accession: B42351
A:Molecule type: protein
A:Residues: 165-178 <ROV2>
A>Note: sequence extracted from NCBI backbone (NCBIP:93851)
A:Accession: C42351
A:Molecule type: protein
A:Residues: 317-333 <ROV3>
A>Note: sequence extracted from NCBI backbone (NCBIP:93856)
A>Note: peptide sequences were determined from rat cDNA cloned and expressed in yeast
C:Superfamily: antithrombin III
C:Keywords: glycoprotein; serine proteinase inhibitor
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-397/Product: proteinase inhibitor nexin I #status predicted <NAT>
F:159/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 2.6%; Score 10; DB 2; Length 397;
Best Local Similarity 100.0%; Pred. No. 0.076;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 VLVNAVYFKG 179
|||||
Db 182 VLVNAVYFKG 191

RESULT 6
A37274

A:Gene: FN-1
C:Superfamily: antithrombin III
C:Keywords: serine proteinase inhibitor
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-397/Product: proteinase inhibitor nexin I #status predicted <NAT>

Glia-derived nexin I precursor, splice form beta - human
N:Alternate names: glia-derived neurite promoting factor; proteinase inhibitor 7; pr
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1991 #sequence revision 31-Dec-1991 #text_change 20-Jun-2000
C:Accession: A37274; J00010; A27496; A26061; A24051
R:McGrogan, M.; Kennedy, J.; Li, M.P.; Hsu, C.; Scott, R.W.; Simonsen, C.C.; Baker, A
BioTechnology 6, 172-177, 1988
A:Title: Molecular cloning and expression of two forms of human protease nexin I.
A:Reference number: J00010
A:Accession: A37274
A:Molecule type: mRNA
A:Residues: 1-398 <MCG1>
A:Experimental source: splice form beta
A:Accession: J00010
A:Molecule type: mRNA
A:Residues: 1-328, 'R', 331-398 <MCG2>
A:Experimental source: splice form alpha
R:Sommer, J.; Gloor, S.M.; Rovelli, G.F.; Hofsteenge, J.; Nick, H.; Meier, R.; Monard
Biochemistry 26, 6407-6410, 1987
A:Title: cDNA sequence coding for a rat glia-derived nexin and its homology to membe
A:Reference number: A90519; MUID:88107544; PMID:3427015
A:Accession: A27496
A:Molecule type: mRNA
A:Residues: 1-328, 'R', 331-398 <SOM>
A:Cross-references: GB:M17783; NID:G183063; PIDN:AAA35883.1; PID:G183064
R:Gloor, S.; Odink, K.; Guenther, J.; Nick, H.; Monard, D.
Cell 47, 687-693, 1986
A:Title: A glia-derived neurite promoting factor with protease inhibitory activity b
A:Reference number: A26061; MUID:87031740; PMID:2877744
A:Accession: A26061
A:Molecule type: protein
A:Residues: 1-259, 'S', 261-398 <GLO>
R:Scott, R.W.; Bergman, B.L.; Baijail, A.; Herish, R.T.; Rodriguez, H.; Jones, B.N.; B
J. Biol. Chem. 260, 7029-7034, 1985
A:Title: Protease nexin. Properties and a modified purification procedure.
A:Reference number: A24051; MUID:85207723; PMID:3997857
A:Accession: A24051
A:Molecule type: protein
A:Residues: 20-47 <SCO>
C:Genetics:
A:Gene: GDB:PI7; PNI
A:Cross-references: GDB:378380; OMIM:177010
A:Map position: 2q33-2q35
C:Superfamily: antithrombin III
C:Keywords: alternative splicing; glycoprotein; serine proteinase inhibitor
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-398/Product: glia-derived nexin I, splice form beta #status experimental <NATA>
F:20-328, 'R', 331-398/Product: glia-derived nexin I, splice form alpha #status predict
F:118,159,394/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:365/Inhibitory site: Arg (thrombin, urokinase) #status predicted

Query Match 2.6%; Score 10; DB 2; Length 398;
Best Local Similarity 100.0%; Pred. No. 0.076;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 VLVNAVYFKG 179
|||||
Db 182 VLVNAVYFKG 191

RESULT 7
A26423
serine proteinase inhibitor 2.1 - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 19-Nov-1988 #sequence_revision 12-Jul-1996 #text_change 12-Jul-1996
C:Accession: A26423
R:Fill, R.B.; Hastie, N.D.
Nature 326, 96-99, 1987
A:Title: Accelerated evolution in the reactive centre regions of serine protease inhi
A:Reference number: A26423; MUID:87144617; PMID:3493437
A:Accession: A26423
A:Molecule type: mRNA

A;Residues: 1-213 <HIL>
 A;Cross-references: GB:X13149
 A;Note: the authors translated the codon CTG for residue 1 as Met, GTG for residue 35 as Met
 C;Superfamily: antithrombin III

Query Match 2.4%; Score 9; DB 2; Length 213;
 Best Local Similarity 100.0%; Pred. No. 0.46;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 333 GTEATAAG 341
 |||||
 Db 165 GTEATAAG 173

RESULT 8
 DXCH
 ovalbumin-related X protein - chicken (fragment)
 C;Species: Gallus gallus (chicken)
 C;Date: 31-Dec-1980 #sequence_revision 31-Dec-1980 #text_change 18-Jun-1999
 C;Accession: A01243
 R;Heilig, R.; Perrin, F.; Gannon, F.; Mardel, J.L.; Chambon, P.
 Cell 20, 625-637, 1980
 A;Title: The ovalbumin gene family: structure of the X gene and evolution of duplicated F13/Binding site: carbohydrate (Asn) (covalent) #status predicted
 A;Reference number: A01243; MUID:81022623; PMID:7418002
 A;Contents: exons 6, 7, and 8
 A;Accession: A01243
 A;Molecule type: DNA
 A;Residues: 1-232 <HEI>
 A;Cross-references: GB:J00920; NID:9212895; PIDN:AAA68881.1; PID:9212897
 C;Superfamily: antithrombin III
 C;Keywords: glycoprotein; phosphoprotein; serine proteinase inhibitor
 F13/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F189/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 2.4%; Score 9; DB 1; Length 233;
 Best Local Similarity 100.0%; Pred. No. 0.5;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 355 FRADHPFLF 363
 |||||
 Db 205 FRADHPFLF 213

RESULT 9
 AC1919
 hypothetical protein all0902 (imported) - Nostoc sp. (strain PCC 7120)
 C;Species: Nostoc sp. PCC 7120
 A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
 C;Accession: AC1919
 R;Kaneke, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shingo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
 A;Reference number: AB1807; MUID:21595285; PMID:11759840
 A;Accession: AC1919
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-248 <KUR>
 A;Cross-references: GB:BA000019; PIDN:BA072859.1; PID:gl7130248; GSPDB:GN00179
 A;Experimental source: strain PCC 7120
 C;Genetics:
 A;Gene: all0902

Query Match 2.4%; Score 9; DB 2; Length 248;
 Best Local Similarity 100.0%; Pred. No. 0.53;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 245 DLSEIENKL 253
 |||||
 Db 236 DLSEIENKL 244

RESULT 10
 M11433
 ovalbumin - Japanese quail
 C;Species: Coturnix coturnix japonica (Japanese quail)
 C;Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999
 C;Accession: M11433
 R;Nucha, J.; Klaudiny, J.; Klaudinyova, V.; Hanes, J.; Simuth, J.
 Nucleic Acids Res. 18, 5553, 1990
 A;Title: The sequence of Japanese quail ovalbumin cDNA.
 A;Reference number: M11433; MUID:91016850; PMID:2216734
 A;Accession: M11433
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-383 <MUC>
 A;Cross-references: EMBL:X53964; NID:962643; PIDN:CAA37916.1; PID:g62644
 C;Superfamily: antithrombin III

Query Match 2.4%; Score 9; DB 2; Length 383;
 Best Local Similarity 100.0%; Pred. No. 0.8;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 355 FRADHPFLF 363
 |||||
 Db 356 FRADHPFLF 364

RESULT 11
 CACH
 ovalbumin [validated] - chicken
 C;Species: Gallus gallus (chicken)
 C;Date: 31-Dec-1979 #sequence_revision 30-Jun-1993 #text_change 15-Sep-2000
 C;Accession: A90455; MUID:150402; A93197; A93827; A90093; A90092; A61297; A42793;
 R;Woo, S.L.C.; Beattie, W.G.; Catterall, J.F.; Dugaiczky, A.; Staden, R.; Brownlee, G.
 Biochemistry 20, 6437-6446, 1981
 A;Title: Complete nucleotide sequence of the chicken chromosomal ovalbumin gene and its expression in the chicken oviduct
 A;Reference number: A90455; MUID:82069038; PMID:6272839
 A;Accession: A90455
 A;Molecule type: DNA
 A;Residues: 1-386 <WOO>
 A;Cross-references: EMBL:V00438; NID:963719; PIDN:CAA23716.1; PID:9808974
 A;Note: a number of silent polymorphic sites are identified and discussed
 A;Note: Thr-188 is also predicted
 R;Catterall, J.F.; O'Malley, B.W.; Robertson, M.A.; Staden, R.; Tanaka, Y.; Brownlee, N.
 Nature 275, 510-513, 1978
 A;Title: Nucleotide sequence homology at 12 intron-exon junctions in the chick ovalbumin gene
 A;Reference number: M15042; MUID:79010682; PMID:692731
 A;Accession: I50402
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-386 <CAT>
 A;Cross-references: GB:M34352; NID:9212501; PIDN:AAA48998.1; PID:g212503
 R;Robertson, M.A.; Staden, R.; Tanaka, Y.; Catterall, J.F.; O'Malley, B.W.; Brownlee, N.
 Nature 278, 370-372, 1979
 A;Title: Sequence of three introns in the chick ovalbumin gene.
 A;Reference number: I50605; MUID:79135070; PMID:423993
 A;Accession: I50605
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-4, 'A', 'F', '120-155 <ROB>
 A;Cross-references: EMBL:V00342; NID:g63051; PIDN:CAA23681.1; PID:g63052
 R;McReynolds, L.; O'Malley, B.W.; Nisbet, A.D.; Pothergill, J.E.; Givol, D.; Fields, N.
 Nature 273, 723-728, 1978
 A;Title: Sequence of chicken ovalbumin mRNA.
 A;Reference number: A93197; MUID:78199842; PMID:661981
 A;Accession: A93197
 A;Molecule type: mRNA
 A;Residues: 1-386 <MCR>
 A;Cross-references: EMBL:V00383; NID:g63053
 A;Note: a minor component has Asp-312
 R;Palmiter, R.D.; Gagnon, J.; Walsh, K.A.
 Proc. Natl. Acad. Sci. U.S.A. 75, 94-98, 1978
 A;Title: Ovalbumin: a secreted protein without a transient hydrophobic leader sequence

A;Reference number: A93827; MUID:78116057; PMID:272676
A;Accession: A93827
A;Molecule type: protein
A;Residues: 2-33, X, 35-36 <PAL>
R;Thompson, E.O.P.; Fisher, W.K.
Aust. J. Biol. Sci. 31, 443-446, 1978
A;Title: A correction and extension of the acetylated amino terminal sequence of ovalbumin
A;Reference number: A90093; MUID:79186958; PMID:751625
A;Accession: A90093
A;Molecule type: protein
A;Residues: 2-17 <TH1>
R;Thompson, E.O.P.; Fisher, W.K.
Aust. J. Biol. Sci. 31, 433-442, 1978
A;Title: Amino acid sequences containing half-cysteine residues in ovalbumin.
A;Reference number: A90092; MUID:79186957; PMID:751624
A;Accession: A90092
A;Molecule type: protein
A;Residues: 6-17,30-36;61-79,116-124,367-374,380-386 <TH2>
R;Tsunasawa, S.; Narita, K.
J. Biochem. 92, 607-613, 1982
A;Title: Micro-identification of amino-terminal acetyl amino acids in proteins.
A;Reference number: A61297; MUID:83056735; PMID:6754709
A;Accession: A61297
A;Molecule type: protein
A;Residues: 2-6 <TSU>
R;Takahashi, N.; Hirose, M.
J. Biol. Chem. 267, 11565-11572, 1992
A;Title: Reversible denaturation of disulfide-reduced ovalbumin and its reoxidation gene
A;Reference number: A42793; MUID:92283876; PMID:1597494
A;Accession: A42793
A;Molecule type: protein
A;Residues: 60-73, X, 75-85,112-119, 'EX', 122-123 <TAK>
R;Stein, P.B.; Leslie, A.G.W.
submitted to the Brookhaven Protein Data Bank, November 1990
A;Reference number: A50294; PDB:1OXA
A;Contents: annotation; X-ray crystallography, 1.95 angstroms, residues 2-386
R;Stein, P.B.; Leslie, A.G.W.; Finch, J.T.; Carrelli, R.W.
J. Mol. Biol. 221, 941-959, 1991
A;Title: Crystal structure of uncleaved ovalbumin at 1.95 Angstroms resolution.
A;Reference number: A58761; MUID:92046044; PMID:1942038
A;Contents: annotation; X-ray crystallography, 1.95 angstroms
C;Genetics:
A;Introns: 56/3; 73/3; 116/3; 156/1; 203/3; 255/3
C;Superfamily: antithrombin III
C;Keywords: acetylated amino end; glycoprotein; phosphoprotein
F;2-364/Products: ovalbumin #status experimental <MAR>
F;2/Modified site: acetylated amino end (Gly) (in mature form) #status experimental
F;89,945/Binding site: phosphate (Ser) (covalent) #status experimental
F;74-121/Disulfide bonds: #status experimental
F;293/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 2.4%; Score 9; DB 1; Length 386;
Best Local Similarity 100.0%; Pred. No. 0.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 355 FRADHPFLF 363
|||||
Db 359 FRADHPFLF 367

RESULT 12
DYCH
ovalbumin-related Y protein - chicken
C;Species: Gallus gallus (chicken)
C;Date: 17-Dec-1982 #sequence_revision 17-Dec-1982 #text_change 18-Jun-1999
C;Accession: A01244
R;Heilig, R.; Muraskovsky, R.; Kloepper, C.; Mandel, J.L.
Nucleic Acids Res. 10, 4363-4382, 1982
A;Title: The ovalbumin gene family: complete sequence and structure of the Y gene.
A;Reference number: A01244; MUID:83014529; PMID:1722240
A;Accession: A01244
A;Molecule type: DNA
A;Residues: 1-388 <HEI>

A;Cross-references: GB:J00922; GB:V00439; NID:g212899; PIDN:AAA68882.1; PID:g212900
C;Genetics:
A;Introns: 56/3; 73/3; 116/3; 156/1; 203/3; 255/3
C;Superfamily: antithrombin III
C;Keywords: glycoprotein; phosphoprotein; serine proteinase inhibitor
F;74-121/Disulfide bonds: #status predicted
F;293/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;7345/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 2.4%; Score 9; DB 1; Length 388;
Best Local Similarity 100.0%; Pred. No. 0.81;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 355 FRADHPFLF 363
|||||
Db 361 FRADHPFLF 369

RESULT 13
S08102
serine proteinase inhibitor 1 - rat
N;Alternate names: growth hormone-induced proteinase inhibitor
C;Species: Rattus norvegicus (Norway rat)
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 16-Jul-1999
C;Accession: S08102; S11318; A29131
R;le Cam, A.
submitted to the EMBL Data Library, August 1989
A;Reference number: S08099
A;Accession: S08102
A;Molecule type: mRNA
A;Residues: 1-403 <LEC>
A;Cross-references: EMBL:X16357; NID:g57230; PIDN:CAA34406.1; PID:g57231
R;Pages, G.; Rouayrenc, J.F.; le Cam, G.; Mariller, M.; le Cam, A.
Eur. J. Biochem. 190, 385-391, 1990
A;Title: Molecular characterization of three rat liver serine-protease inhibitors af
A;Reference number: S11318; MUID:90306038; PMID:1694763
A;Accession: S11318
A;Status: Preliminary
A;Molecule type: mRNA
A;Residues: 1-403 <PAG>
A;Cross-references: EMBL:X16357; NID:g57230; PIDN:CAA34406.1; PID:g57231
R;Yoon, J.B.; Towle, H.C.; Seelig, S.
J. Biol. Chem. 262, 4284-4289, 1987
A;Title: Growth hormone induces two mRNA species of the serine protease inhibitor ge
A;Reference number: A92632; MUID:87166046; PMID:3494016
A;Accession: A29131
A;Molecule type: mRNA
A;Residues: 82-234, 'L', 236-403 <YOO>
A;Cross-references: GB:M15917; GB:J02692; NID:g207041; PIDN:AAA42172.1; PID:g207042
C;Superfamily: antithrombin III

Query Match 2.4%; Score 9; DB 2; Length 403;
Best Local Similarity 100.0%; Pred. No. 0.84;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 333 GTENTATG 341
|||||
Db 355 GTENTATG 363

RESULT 14
T31297
hypothetical protein - Sphingomonas aromaticivorans plasmid pNL1
C;Species: Sphingomonas aromaticivorans
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-Jan-2000
C;Accession: T31297
R;Romine, M.F.; Stillwell, L.C.; Wong, K.K.; Thurston, S.J.; Sisk, E.C.; Sensen, C.W.
submitted to the EMBL Data Library, July 1998
A;Description: Complete sequence of a 184 kb catabolic plasmid from Sphingomonas aro
A;Reference number: Z20992
A;Accession: T31297
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA

A:Residues: 1-121 <ROM>
 A:Cross-references: EMBL:AF079317; NID:G3378261; PID:G3378438; PIDN:AADC4021.1
 C:Genetics:
 A:Genome: plasmid pNL1
 A>Note: pchC

Query Match 2.1%; Score 8; DB 2; Length 121;
 Best Local Similarity 100.0%; Pred. No. 3;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 LSLFAALA 40
 |||||
 Db 7 LSLFAALA 14

RESULT 15

B8327
 hypothetical protein BHI418 [imported] - Bacillus halodurans (strain C-125)
 C:Species: Bacillus halodurans
 C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 C:Accession: B83827
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
 Nucleic Acids Res. 28, 4317-4331, 2000
 A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
 A:Reference number: A83650; MUID:20512582; PMID:11058132
 A:Accession: B83827
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-262 <STO>
 A:Cross-references: GB:AP001512; GB:BA000004; NID:G10174030; PIDN:BA0C5137.1; GSPDB:GNCO
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: BHI418
 C:Superfamily: hypothetical protein M0933

Query Match 2.1%; Score 8; DB 2; Length 262;
 Best Local Similarity 100.0%; Pred. No. 6.1;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ASLAANA 9
 |||||
 Db 111 ASLAANA 118

Search completed: December 12, 2003, 16:35:17
 Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 16:31:13 ; Search time 40 Seconds
(without alignments)
2451.500 Million cell updates/sec

Title: US-09-508-997A-2

Perfect score: 380

Sequence: 1 MASLAANAEPFCNLFREMD.....FVIRKDDIILSGVSCP 380

Scoring table: OLIGO

Gapop 60.0, Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size: 0

Total number of hits satisfying chosen parameters: 930525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp Vertebrate:*
- 14: sp Unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteria:*
- 17: sp_archae:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	18	4.7	380	11 Q9D695	Q9d695 mus musculus
2	18	4.7	380	11 Q9D695	Q9d695 rattus norv
3	14	3.7	334	12 Q9DHG4	Q9dhg4 yaba-like d
4	11	2.9	385	11 Q8VHQ0	Q8vhq0 mus musculus
5	11	2.9	388	11 Q9Z2G2	Q9z2g2 mus musculus
6	10	2.6	292	6 Q8WNWA	Q8wnwa sus scrofa
7	10	2.6	359	11 Q9D6A7	Q9d6a7 mus musculus
8	10	2.6	366	11 Q9DLE7	Q9dle7 mus musculus
9	10	2.6	387	11 Q9DIQ5	Q9diq5 mus musculus
10	10	2.6	387	11 Q8BHL1	Q8bhl1 mus musculus
11	10	2.6	387	11 Q8BGR6	Q8bgr6 mus musculus
12	10	2.6	397	6 Q8HZX1	Q8hzy1 bos taurus
13	9	2.4	248	16 Q8VYE8	Q8vye8 anabaena sp
14	9	2.4	423	11 Q9DFP9	Q9dfp9 mus musculus
15	9	2.4	454	16 Q92K34	Q92k34 rhizobium m
16	9	2.4	615	4 Q8TAP0	Q8tap0 homo sapien

Q9mrx3 homo sapien
Q86005 springomona
Q96dw8 homo sapien
Q8bnw4 mus musculus
Q8ixh2 homo sapien
Q9kd01 bacillus ha
Q9gz39 trypanosoma
Q97ig4 clostridium
Q02377 manduca sex
Q97h35 clostridium
Q92qu5 rhizobium m
Q9und9 homo sapien
Q8nl77 homo sapien
Q9va64 drosophila
Q9x7y5 streptomyce
Q9a2k3 caulobacter
Q92kk6 rhizobium m
Q8pvi7 methanosarc
Q9v2a8 pyrobaculum
Q8zwl8 pyrobaculum
Q59462 pyrococcus
Q95z82 leishmania
Q65514 escherichia
Q9s8n3 triticum ae
Q8kbf2 chlorobium
Q8h8t3 oryza sativ
Q01961 bacterioph
Q8v038 anabaena sp
Q9lun6 rhizobium m

ALIGNMENTS

RESULT 1

Q9D695
ID Q9D695 PRELIMINARY; PRT; 380 AA.
AC Q9D695;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE 4631416M05Rik protein (Megin).
GN SERPINB7 OR 431416M05RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CS7BU/60; TISSUE=Skin;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staehli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baidarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Buit C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.

```

RX MEDLINE=21368006; PubMed=11473647;
RA Nangaku M., Miyata T., Suzuki D., Umezono T., Hashimoto T., Wada T.,
RT Yagi M., Nagano N., Inagi R., Kurokawa K.;
RT "Cloning of rodent mesgin revealed its up-regulation in
RT mesangioproliferative nephritis.";
RL Kidney Int. 60:641-652(2001).
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL; AK014524; BA329410.1; -.
DR EMBL; AF105328; AAL16768.1; -.
DR HSSP; P05121; 1A7C.
DR MGD; MG1:2151053; Serpinb7.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; serpin; 1.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
KW Protease inhibitor; Serine protease inhibitor; Serpin.
SQ SEQUENCE 380 AA; 43050 MW; C9240272BCFB9CF4 CRC64;

Query Match 4.7%; Score 18; DB 11; Length 380;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 SSSAVWVLNNAVYFKGW 181
Db 164 SSSAVWVLNNAVYFKGW 181

RESULT 2
Q920J5 PRELIMINARY; PRT; 380 AA.
AC Q920J5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Mesgin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21368006; PubMed=11473647;
RA Nangaku M., Miyata T., Suzuki D., Umezono T., Hashimoto T., Wada T.,
RT Yagi M., Nagano N., Inagi R., Kurokawa K.;
RT "Cloning of rodent mesgin revealed its up-regulation in
RT mesangioproliferative nephritis.";
RL Kidney Int. 60:641-652(2001).
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL; AF105329; AAL16769.1; -.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; serpin; 1.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
KW Protease inhibitor; Serine protease inhibitor; Serpin.
SQ SEQUENCE 380 AA; 42821 MW; D8076CABEE2C2FBC CRC64;

Query Match 4.7%; Score 13; DB 11; Length 380;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 SSSAVWVLNNAVYFKGW 181
Db 164 SSSAVWVLNNAVYFKGW 181

RESULT 3
Q9DHG4 PRELIMINARY; PRT; 334 AA.
AC Q9DHG4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE 149R protein.

```

```

GN 149R.
OS Yaba-like disease virus (YLDV).
OC Viruses; dsRNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Yatapoxvirus.
OX NCBI_TaxID=132475;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee H.J.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21176366; PubMed=11277691;
RA Lee H.J., Essani K., Smith G.L.;
RT "The genome sequence of Yaba-like disease virus, a yatapoxvirus.";
RL Virology 281:170-192(2001).
RN [3]
RP SEQUENCE FROM N.A.
RA Lee H.J.;
RL Thesis (2000), Sir William Dunn School of Pathology, University of
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL; AJ293568; CAC21387.1; -.
DR HSSP; P05120; 1BY7.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; serpin; 1.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
KW Protease inhibitor; Serine protease inhibitor; Serpin.
SQ SEQUENCE 334 AA; 38190 MW; D7FD004191B3C19F CRC64;

Query Match 3.7%; Score 14; DB 12; Length 334;
Best Local Similarity 100.0%; Pred. No. 7.1e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 323 HKSYIEVTEEGTEA 336
Db 275 HKSYIEVTEEGTEA 288

RESULT 4
Q8VHQ0 PRELIMINARY; PRT; 385 AA.
AC Q8VHQ0;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE SP31L2.
GN C76171 OR SP31L2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Kalseman D., Knaggs S., Scarff K.L., Gillard A., Mirza G., Cadman M.,
RA McKeone R., Denny P., Cooley J., Benarafa C., Remold-O'Donnell E.,
RA Ragoussis J., Bird P.I.;
RT "Comparison of Human Chromosome 6p25 with Murine Chromosome 13 Reveals
RT a Greatly Expanded Oxy-Serpin Gene Repertoire in the Mouse.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL; AF425084; AAL65910.1; -.
DR HSSP; P01009; 1QLP.
DR MGD; MG1:2145481; C76171.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; serpin; 1.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
KW Protease inhibitor; Serine protease inhibitor; Serpin.
SQ SEQUENCE 385 AA; 43786 MW; 0820A9A26AEBB485 CRC64;

Query Match 2.9%; Score 11; DB 11; Length 385;
Best Local Similarity 100.0%; Pred. No. 0.013;

```

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 171 LVNAVYFKGW 181
 Db 165 LVNAVYFKGW 175

RESULT 5

Q922G2
 ID Q922G2 PRELIMINARY; PRT; 388 AA.
 AC Q922G2;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Squamous cell carcinoma antigen 2.
 GN SERPINB4 OR SCCA2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 [1]_
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99047536; PubMed=3828132;
 RA Baruski A.J., Kamachi Y., Schick C., Massa H., Trask B.J.,
 RA Silverman G.A.;
 RT "A murine ortholog of the human serpin SCCA2 maps to chromosome 1 and
 RT inhibits chymotrypsin-like serine proteinases."
 RL Genomics 54:297-306(1998).
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
 DR EMBL; AF063937; AAC95432.1; -;
 DR HSSP; P05120; 1BY7.
 DR MGD; MGI:1277952; Serpinb4.
 DR InterPro; IPR000215; Serpin.
 DR Pfam; PF00079; serpin; 1.
 DR SMART; SM00093; SERPIN; 1.
 DR PROSITE; PS00284; SERPIN; 1.
 KW Protease inhibitor; Serine protease inhibitor; Serpin.
 SQ SEQUENCE 388 AA; 44573 MW; D77AFA1233EB9CD CRC64;

Query Match 2.9%; Score 11; DB 11; Length 388;
 Best Local Similarity 100.0%; Pred. No. 0.013;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 MVLNAVYFKG 179
 Db 172 MVLNAVYFKG 182

RESULT 6

Q8WNW8
 ID Q8WNW8 PRELIMINARY; PRT; 292 AA.
 AC Q8WNW8;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Nexin-1 (Fragment).
 GN PN-1.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 [1]_
 RN
 RP SEQUENCE FROM N.A.
 RA Paradis V., Lussier J., Silversides D.W.;
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
 DR EMBL; AY059415; AAL23838.2; -;
 DR InterPro; IPR000215; Serpin.
 DR Pfam; PF00079; serpin; 1.
 DR SMART; SM00093; SERPIN; 1.
 DR PROSITE; PS00284; SERPIN; 1.
 KW Protease inhibitor; Serine protease inhibitor; Serpin.
 FT NON_TER 1 1

SQ SEQUENCE 292 AA; 32597 MW; 350ACFF1DF6DC502 CRC64;

Query Match 2.6%; Score 10; DB 6; Length 292;
 Best Local Similarity 100.0%; Pred. No. 0.12;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 VLVNAVYFKG 179

Db 77 VLVNAVYFKG 86

RESULT 7

Q9D6A7
 ID Q9D6A7 PRELIMINARY; PRT; 359 AA.
 AC Q9D6A7;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE 393042J058Eik protein.
 GN SP111 OR 393042J058EIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 [1]_
 RN
 RP SEQUENCE FROM N.A.
 RX STRAIN=C57BL/6J; TISSUE=Placenta, and Extraembryonic tissue;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai T., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Aizawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamahata I.,
 RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli P., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barish G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang X.H., Weiss C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-693(2001).
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
 DR EMBL; AK014448; BAB29358.1; -;
 DR HSSP; P05120; 1BY7.
 DR MGD; MGI:894669; Sp11.
 DR InterPro; IPR000215; Serpin.
 DR Pfam; PF00079; serpin; 1.
 DR SMART; SM00093; SERPIN; 1.
 DR PROSITE; PS00284; SERPIN; 1.
 KW Protease inhibitor; Serine protease inhibitor; Serpin.
 SQ SEQUENCE 359 AA; 41091 MW; C1109BE74A4814BD CRC64;

Query Match 2.6%; Score 10; DB 11; Length 359;
 Best Local Similarity 100.0%; Pred. No. 0.14;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 331 ECGTEAAT 340

Db 307 ECGTEAAT 316

RESULT 8

Q9D1E7
 ID Q9D1E7 PRELIMINARY; PRT; 386 AA.
 AC Q9D1E7;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)

RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA	Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA	Fleischman W., Gaasterland I., Gissi C., King B., Kochiwa H.,
RA	Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA	Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA	Sakai K., Okido T., Furuno M., Aono H., Balidarelli R., Barè G.,
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA	Gustincich S., Hill D., Hofmann W., Hume D.A., Kamiya M., Lee N.H.,
RA	Lions P., Marchionni L., Masima J., Mazzarello J., Momabaerts P.,
RA	Nordone P., Ring B., Schoenwald M., Rodriguez I., Sakamoto N.,
RA	Sasaki H., Sato K., Schönbach C., Seya T., Shibata Y., Storch K.-F.,
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtesuki S.,
RA	Hayashizaki Y.,
RA	"Functional annotation of a full-length mouse cDNA collection.";
RT	Nature 409:685-690(2001).
CC	- - SIMILARITY: BELONGS TO THE SERPIN FAMILY.
CC	EMBL; AK003220; BAB2650.1; -.
DR	HSP; P05120; 1BY7.
DR	IiterPro; IPR000215; Serpin.
DR	Pfam; PF000079; serpin; 1.
DR	SMART; PS000093; SERPIN; 1.
DR	PROSITE; PS00284; SERPIN; 1.
KW	Protease inhibitor; Serine protease inhibitor; Serpin.
SQ	SEQUENCE 387 AA; 44558 MW; AE25C02503FB194 CRC64;
Query Match	2.6%; Score 10; DB 11; Length 387;
Best Local Similarity	100.0%; Pred. No. 0.15; Indels 0; Gaps 0;
Matches	10; Conservative 0; Mismatches 0;
QY	170 VLVNAVVFVKG 179
DB	172 VLVNAVVFVKG 181
RESULT 10	
QB5HL1	
ID	Q8BHL1 PRELIMINARY; PRT; 387 AA.
AC	Q8BHL1;
DT	01-MAR-2003 (TrEMBLrel. 23, Created)
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE	Squamous cell carcinoma antigen 2 related protein 1.
GN	SCCA2-RS1.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=Car-S, and Car-R; TISSUE=TPA-treated skin;
RA	Gariboldi M., Peissel B., Fabri A., Saran A., Zaffaroni D.,
RA	Spinola M., Falvella P.S., Pazzaglia S., Tanuma J.-I., Maurichi A.,
RA	Bartoli C., Silverman G., Covelli V., Pilotti S., Hayashizaki Y.,
RA	Okazaki Y., Dragani T.A.,
RT	"The serpin Scca2 gene plays a functional role in genetic
RT	susceptibility to skin tumorigenesis in mice and humans."
RL	submitted (Aug-2002) to the EMBL/GenBank/DBJ databases."
DR	EMBL; AY144685; AAN62872.1; -.
DR	EMBL; AY144685; AAN62873.1; -.
SQ	SEQUENCE 387 AA; 44539 MW; 0AD6396FA0251135 CRC64;
Query Match	2.6%; Score 10; DB 11; Length 387;
Best Local Similarity	100.0%; Pred. No. 0.15;
Matches	10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	170 VLVNAVVFVKG 179
DB	172 VLVNAVVFVKG 181
RESULT 11	

```

Q8B956
ID Q8B956 PRELIMINARY; PRT; 387 AA.
AC Q8B956;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Squamous cell carcinoma antigen 2.
GN SERPINB4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STEAN-Gar-S, and Car-P; TISSUE=TPA-treated skin;
RA Gariboldi M., Peissel B., Fabbri A., Saran A., Zaffaroni D.,
RA Spinola M., Falvella F.S., Pazzaglia S., Tanuma J.-I., Maurichi A.,
RA Bartoli C., Silverman G., Coveili V., Pilotti S., Hayashizaki Y.,
RA Okazaki Y., Dragani T.A.;
RT "The serpin Scga2 gene plays a functional role in genetic
RT susceptibility to skin tumorigenesis in mice and humans.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY144683; AAN62870.1; -.
DR EMBL; AY144684; AAN62871.1; -.
SQ SEQUENCE 387 AA; 44434 MW; B5B15F807FEF0688 CRC64;

Query Match 2.6%; Score 10; DB 11; Length 387;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 VLNVAVYFKG 179
Db 172 VLNVAVYFKG 181
|||||

RESULT 12
Q8HZY1
ID Q8HZY1 PRELIMINARY; PRT; 397 AA.
AC Q8HZY1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Serine protease inhibitor-E2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC Bedard J., Brule S., Price C., Silversides D.W., Lussier J.;
RT "Serine protease inhibitor-E2 (SERPINE2) is differentially expressed
RT in granulosa cells of dominant follicle in cattle.";
RL Mol. Reprod. Dev. 0:0-0(2002).
DR EMBL; AF251153; AAN37922.1; -.
KW Protease.
SQ SEQUENCE 397 AA; 43877 MW; 90893C65D6A15443 CRC64;

Query Match 2.6%; Score 10; DB 6; Length 397;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 VLNVAVYFKG 179
Db 182 VLNVAVYFKG 191
|||||

RESULT 13
Q8YF58
ID Q8YF58 PRELIMINARY; PRT; 246 AA.
AC Q8YF58;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

```

```

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein AL10902.
GN AL10902.
OS Arabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2159285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iwaguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Arabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003584; BAB72859.1; -.
DR InterPro; IPR001853; DSEA.
DR InterPro; IPR006663; ThioRedox_dom2.
DR Pfam; PF01323; DSEA; 1.
DR Hypothetical protein; Complete proteome.
KW SEQUENCE 248 AA; 27809 MW; A5BFF1919A66E2DA CRC64;

Query Match 2.4%; Score 9; DB 16; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 245 DLSEIENK 253
Db 236 DLSEIENK 244
|||||

RESULT 14
Q9D7P9
ID Q9D7P9 PRELIMINARY; PRT; 423 AA.
AC Q9D7P9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE 2300003F07RIK protein (Weakly similar to serpin B12).
GN 2300003F07RIK
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann M., Gaasterland T., Gissi C., King S., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schiraldi L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.P.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momberti P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sakaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kontuski S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;

```



```

RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RT the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL; AK09018; BAB26028.1; -.
DR EMBL; AK040697; BAC30672.1; -.
DR HSP; P05619; 1HLE.
DR MSD; MGI:1919119; 2300003F07Rik.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; serpin; 1.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
KW Protease inhibitor; Serine protease inhibitor; Serpin.
SQ SEQUENCE 423 AA; 47834 MW; 5A22BE2FE51B6120 CRC64;

Query Match 2.4%; Score 9; DB 11; Length 423;
Best Local Similarity 100.0%; Pred.No.1.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 VLVNAVYFK 178
DB 204 VLVNAVYFK 212
|||||

RESULT 15
Q92K34
ID Q92K34 PRELIMINARY; PRT; 454 AA.
AC Q92K34;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Hypothetical protein R02208.
GN R02208 OR SMC01575.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Rubier F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puhler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
DR EMBL; AL591790; CAC45787.1; -.
DR InterPro; IPR005490; BxfK Ybis YhmG.
DR Pfam; PF03734; BxfK Ybis YhmG; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 454 AA; 49052 MW; 6C25A2A60730763D CRC64;

Query Match 2.4%; Score 9; DB 16; Length 454;
Best Local Similarity 100.0%; Pred.No.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 LSLFAALAL 41
DB 9 LSLFAALAL 17
|||||

```

Search completed: December 12, 2003, 16:34:53
Job time : 52 secs

Result	Query	Score	Match	Length	ID	Description
1	1959	100.0	380	15	AAR48379	Human megakaryocyt
2	1959	100.0	380	20	AAU08254	Human megsin prote
3	1959	100.0	380	21	AAB24142	Human megsin prote
4	1959	100.0	380	22	AAB30075	Human megsin prote
5	1959	100.0	380	24	ABU56533	Human megsin prote
6	1488	76.0	380	22	AAQ54286	Lung cancer-associ
7	1777	75.4	380	20	AAQ08255	Rat megsin protein
8	1477	75.4	380	21	AAB24150	Rat megsin protein
9	1477	75.4	380	22	AAB30076	Rat megsin protein

PI Ishida N, Iwasa F, Kurihara T, Miura K, Nakazato H;
PI Tsujimoto M, Tsuruoka N, Yamaguchi N, Yamaichi K;
PI

DR WPI; 1994-050782/08.
 DR N-PSDB; AAQ56670.
 XX
 PT New megakaryocyte differentiation factor - isolated from human
 PT epidermoid carcinoma cells, used to treat conditions involving a
 PT decrease in platelets
 XX
 XX Claim 7; Page 30-32; 47pp; English.
 XX
 CC Human MDF can be isolated from a culture of human epidermoid
 CC carcinoma A431 cells in protein-free medium. The MDF stimulates
 CC differentiation of megakaryocytes from myeloid cells in the presence
 CC of IL-3. The MDF acts in vivo as a thrombopoietin making it useful
 CC for treatment of diseases involving a decrease in platelet number
 CC (esp. thrombocytopenia) such as occurs in bone marrow
 CC transplantation and in chemotherapy. MDF has mol.wt. 55-57kD by
 CC SDS-PAGE and exhibits an isoelectric point of 6.0-7.0.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 XX Sequence 380 AA;
 SQ
 Query Match 100.0%; Score 1959; DB 15; Length 380;
 Best Local Similarity 100.0%; Pred. No. 3.4e-169;
 Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MASLAANAARFCFNLFREDDNQNGNVFFSSLSLPAALALVRLGAQDDSLSDIKLLHV 60
 DB 1 MASLAANAARFCFNLFREDDNQNGNVFFSSLSLPAALALVRLGAQDDSLSDIKLLHV 60
 QY 61 NTASGYGSSNSQSGLSQQLKRVFSDINASHKDYDLSIVNGLFAEKYVGFHKDYIECAEK 120
 DB 61 NTASGYGSSNSQSGLSQQLKRVFSDINASHKDYDLSIVNGLFAEKYVGFHKDYIECAEK 120
 QY 121 LYDAKVERVDFTNLEEDTRRNINKWVENETHGKIKRNVIGEGGSSSAVMVLNAVYFKG 180
 DB 121 LYDAKVERVDFTNLEEDTRRNINKWVENETHGKIKRNVIGEGGSSSAVMVLNAVYFKG 180
 QY 181 WQSAFTKETINCHFKSPKCGKAVAMHQRKFNLSVIEDPSMKILELYNGGINMYVL 240
 DB 181 WQSAFTKETINCHFKSPKCGKAVAMHQRKFNLSVIEDPSMKILELYNGGINMYVL 240
 QY 241 LPENDLSIENKLTFFQNLMEWTPRMTSKYVEVFPQPKIEKNYEMKQYLRALGLKDF 300
 DB 241 LPENDLSIENKLTFFQNLMEWTPRMTSKYVEVFPQPKIEKNYEMKQYLRALGLKDF 300
 QY 301 DESKADLSIASGGRLYISRMHKSIVTEEGTEATAATGSNIVEKQLPQSTLFRADHP 360
 DB 301 DESKADLSIASGGRLYISRMHKSIVTEEGTEATAATGSNIVEKQLPQSTLFRADHP 360
 QY 361 FLFVIRKDDIILFSGKVSCP 380
 DB 361 FLFVIRKDDIILFSGKVSCP 380
 RESULT 2
 AAY08254
 ID AAY08254 standard; Protein; 380 AA.
 XX
 AC AAY08254;
 XX
 XX 14-JUL-1999 (first entry)
 DT
 DE Human megsin protein.
 XX
 XX
 KW Megsin; mesangial cell; treatment; diagnosis; disease; IgA nephropathy;
 KW human; rat; murine.
 XX
 XX Homo sapiens.
 OS
 XX
 PN W05915652-A1.
 XX
 XX
 PD 01-APR-1999.
 XX

PF 22-SEP-1998; 98WO-JP04269.
 XX
 PR 22-SEP-1997; 97JP-0275302.
 XX
 PA (KUJO/) KUROKAWA K.
 PA (MIYA/) MIYATA T.
 XX
 PI Miyata T;
 XX
 DR WPI; 1999-276983/23.
 DR N-PSDB; AAX56712.
 XX
 PT Megsin protein expressed specifically in mesangial cells
 XX
 XX Claim 1; Page 62-64; 100pp; Japanese.
 PS
 CC This invention describes the isolation of novel megsin nucleic acid and
 CC proteins from human, rat and mouse tissue. This protein is expressed
 CC specifically in mesangial cells. The products of the invention are
 CC useful for the treatment and diagnosis of diseases involving mesangial
 CC cells, such as IgA nephropathy.
 XX
 XX Sequence 380 AA;
 SQ
 Query Match 100.0%; Score 1959; DB 20; Length 380;
 Best Local Similarity 100.0%; Pred. No. 3.4e-169;
 Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MASLAANAARFCFNLFREDDNQNGNVFFSSLSLPAALALVRLGAQDDSLSDIKLLHV 60
 DB 1 MASLAANAARFCFNLFREDDNQNGNVFFSSLSLPAALALVRLGAQDDSLSDIKLLHV 60
 QY 61 NTASGYGSSNSQSGLSQQLKRVFSDINASHKDYDLSIVNGLFAEKYVGFHKDYIECAEK 120
 DB 61 NTASGYGSSNSQSGLSQQLKRVFSDINASHKDYDLSIVNGLFAEKYVGFHKDYIECAEK 120
 QY 121 LYDAKVERVDFTNLEEDTRRNINKWVENETHGKIKRNVIGEGGSSSAVMVLNAVYFKG 180
 DB 121 LYDAKVERVDFTNLEEDTRRNINKWVENETHGKIKRNVIGEGGSSSAVMVLNAVYFKG 180
 QY 181 WQSAFTKETINCHFKSPKCGKAVAMHQRKFNLSVIEDPSMKILELYNGGINMYVL 240
 DB 181 WQSAFTKETINCHFKSPKCGKAVAMHQRKFNLSVIEDPSMKILELYNGGINMYVL 240
 QY 241 LPENDLSIENKLTFFQNLMEWTPRMTSKYVEVFPQPKIEKNYEMKQYLRALGLKDF 300
 DB 241 LPENDLSIENKLTFFQNLMEWTPRMTSKYVEVFPQPKIEKNYEMKQYLRALGLKDF 300
 QY 301 DESKADLSIASGGRLYISRMHKSIVTEEGTEATAATGSNIVEKQLPQSTLFRADHP 360
 DB 301 DESKADLSIASGGRLYISRMHKSIVTEEGTEATAATGSNIVEKQLPQSTLFRADHP 360
 QY 361 FLFVIRKDDIILFSGKVSCP 380
 DB 361 FLFVIRKDDIILFSGKVSCP 380
 RESULT 3
 AAB24142
 ID AAB24142 standard; Protein; 380 AA.
 XX
 AC AAB24142;
 XX
 XX 30-JAN-2001 (first entry)
 DT
 DE Human megsin protein sequence SEQ ID NO:2.
 XX
 XX
 KW Megsin; mesangium-predominant gene; serpin regulated; nephropathy;
 KW IgA; immunoglobulin A; detection; renal function; renal disorder;
 KW diagnosis; biological sample; blood; urine.
 XX
 XX Homo sapiens.
 CS
 XX

```

PN WC200057189-A1.
XX
XX PD
XX PF
XX PF 17-MAR-2000; 2000WO-JP01646.
XX
XX PR 19-MAR-1999; 99JP-0075305.
XX PR 28-OCT-1999; 99JP-0306623.
XX
XX PA (KURO/) KUROKAWA K.
XX PA (FUSO) FUSO PHARM IND LTD.
XX PA (MIYA/) MIYATA T.
XX
XX PI Miyata T;
XX
XX WI: 2000-611642/58.
XX DR N-PSDB; AAA99294.
XX
XX PT Evaluating renal function comprises assaying mesgin protein in
XX PT biological sample -
XX
XX PS Example 2; Page 66-69; 93pp; Japanese.
XX
XX CC The present invention describes a method for evaluating renal function.
XX CC The method comprises assaying mesgin protein in biological sample. Also
XX CC described are: (1) use of a anti-mesgin protein antibody for diagnosing
XX CC renal function; and (2) a kit for detecting mesgin protein comprising:
XX CC (a) anti-mesgin protein antibody attached to solid magnetic particles;
XX CC (b) direct or indirect fixing for the antibody to the particles; and
XX CC (c) a magnet. The process is useful for evaluating renal function and
XX CC diagnosing renal disorders by assaying mesgin protein in biological
XX CC samples (preferably urine or blood). The process is reproducible and
XX CC gives accurate results. The present sequence represents the human mesgin
XX CC protein, which is given in the exemplification of the present invention.
XX
XX SQ Sequence 380 AA;
XX
XX Query Match 100.0%; Score 1959; DB 21; Length 380;
XX Best Local Similarity 100.0%; Pred. No. 3.4e-169;
XX Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MASLAAANAECFNLFFEMDDNQGNVFFSSLSLFAALALVRLGAQDDSLSDIKLLHV 60
Db 1 MASLAAANAECFNLFFEMDDNQGNVFFSSLSLFAALALVRLGAQDDSLSDIKLLHV 60
QY 61 NTASGVCNSSQSGLSQIKVPSDINASHKDYDLSIVNGLFAEKYVGHKDYIECAEK 120
Db 61 NTASGVCNSSQSGLSQIKVPSDINASHKDYDLSIVNGLFAEKYVGHKDYIECAEK 120
QY 121 LYDAKVERVDFTNLEDTNRNINKWVENETHGKIKNVI GEGGISSAVMVLNNAVYFKGK 180
Db 121 LYDAKVERVDFTNLEDTNRNINKWVENETHGKIKNVI GEGGISSAVMVLNNAVYFKGK 180
QY 181 WQSAFTKSETTINCHFKSPKCSGKAVAMHQRKFNLSVIEDPSMKILRLYNGGINMYVL 240
Db 181 WQSAFTKSETTINCHFKSPKCSGKAVAMHQRKFNLSVIEDPSMKILRLYNGGINMYVL 240
QY 241 LPENDLSIENKLTFFQNLMEWTPRMTSKYVEVFPFPQFKIEKNYKQVLRALGKDIF 300
Db 241 LPENDLSIENKLTFFQNLMEWTPRMTSKYVEVFPFPQFKIEKNYKQVLRALGKDIF 300
QY 301 DESKADLSGIASGRULYISRMWHKSYIEVTEEGTEATATGNSNIVEKQLPOSTLPRADHP 360
Db 301 DESKADLSGIASGRULYISRMWHKSYIEVTEEGTEATATGNSNIVEKQLPOSTLPRADHP 360
QY 361 FLFVIRKDDIILFSGKVSCP 380
Db 361 FLFVIRKDDIILFSGKVSCP 380

```

RESULT 4

AAB83075

ID AAB83075 standard; Protein; 380 AA.

```

XX AAB83075;
XX
XX DT 10-JUL-2001 (first entry)
XX
XX DE Human mesgin protein.
XX
XX KW Human; mesgin; mesangial cell proliferative nephritis; nephrotropic;
XX KW transgenic mouse; glomerular disease; animal model; drug screening.
XX
XX OS Homo sapiens.
XX
XX PN WO200124628-A1.
XX
XX PD 12-APR-2001.
XX
XX PF 06-OCT-2000; 2000WO-JP06988.
XX
XX PR 06-OCT-1999; 99JP-0285736.
XX
XX PA (KURO/) KUROKAWA K.
XX PA (MIYA/) MIYATA T.
XX
XX PI Miyata T;
XX
XX WI: 2001-300136/31.
XX DR N-PSDB; AAF82438.
XX
XX PT Mouse model for mesangial cell proliferative nephritis for development
XX PT and screening of new treatments -
XX
XX PS Example 4; Page 44-46; 62pp; Japanese.
XX
XX CC The present sequence is human mesgin. The human mesgin coding
XX CC sequence may be introduced into a mouse to produce an animal model of
XX CC mesangial cell proliferative nephritis. The symptoms include
XX CC enlargement of the mesangial base region, sedimentation of an immune
XX CC complex and an increase in mesangial cells. The animal model is useful
XX CC for analysing the pathology of chronic glomerular diseases and for
XX CC screening compositions for prevention and treatment of the diseases.
XX CC Highly uniform models can be made easily and in large numbers using
XX CC this method.
XX
XX SQ Sequence 380 AA;

```

```

Query Match 100.0%; Score 1959; DB 22; Length 380;
Best Local Similarity 100.0%; Pred. No. 3.4e-169;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASLAAANAECFNLFFEMDDNQGNVFFSSLSLFAALALVRLGAQDDSLSDIKLLHV 60
Db 1 MASLAAANAECFNLFFEMDDNQGNVFFSSLSLFAALALVRLGAQDDSLSDIKLLHV 60
QY 61 NTASGVCNSSQSGLSQIKVPSDINASHKDYDLSIVNGLFAEKYVGHKDYIECAEK 120
Db 61 NTASGVCNSSQSGLSQIKVPSDINASHKDYDLSIVNGLFAEKYVGHKDYIECAEK 120
QY 121 LYDAKVERVDFTNLEDTNRNINKWVENETHGKIKNVI GEGGISSAVMVLNNAVYFKGK 180
Db 121 LYDAKVERVDFTNLEDTNRNINKWVENETHGKIKNVI GEGGISSAVMVLNNAVYFKGK 180
QY 181 WQSAFTKSETTINCHFKSPKCSGKAVAMHQRKFNLSVIEDPSMKILRLYNGGINMYVL 240
Db 181 WQSAFTKSETTINCHFKSPKCSGKAVAMHQRKFNLSVIEDPSMKILRLYNGGINMYVL 240
QY 241 LPENDLSIENKLTFFQNLMEWTPRMTSKYVEVFPFPQFKIEKNYKQVLRALGKDIF 300
Db 241 LPENDLSIENKLTFFQNLMEWTPRMTSKYVEVFPFPQFKIEKNYKQVLRALGKDIF 300
QY 301 DESKADLSGIASGRULYISRMWHKSYIEVTEEGTEATATGNSNIVEKQLPOSTLPRADHP 360
Db 301 DESKADLSGIASGRULYISRMWHKSYIEVTEEGTEATATGNSNIVEKQLPOSTLPRADHP 360

```

QY 361 FLVIRKDDIILFSGKVSCP 380
 DB 361 FLVIRKDDIILFSGKVSCP 380

RESULT 5

ABU56533
 ID ABU56533 standard; Protein; 380 AA.

XX AC ABU56533;

XX DT 02-APR-2003 (first entry)

XX DE Lung cancer-associated polypeptide #126.

XX KW Lung cancer-associated polypeptide; cytostatic; emphysema;
 KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

XX OS Unidentified.

XX PN WO200286443-A2.

XX PD 31-OCT-2002.

XX PF 18-APR-2002; 2002WO-US12476.

XX PR 18-APR-2001; 2001US-284770P.

XX PR 10-MAY-2001; 2001US-290492P.

XX PR 09-NOV-2001; 2001US-339245P.

XX PR 13-NOV-2001; 2001US-350866P.

XX PR 29-NOV-2001; 2001US-334370P.

XX PR 12-APR-2002; 2002US-372246P.

XX PA (EOSB-) EOS BIOTECHNOLOGY INC.

XX PI Aziz N, Murray R;

XX XX WPI; 2003-093161/08.

DR N-PSDB; ABX76260.

XX PT Detecting a lung cancer-associated transcript in a cell from a patient

XX PT for treating lung cancer, by contacting a biological sample from the

XX PT patient with a polynucleotide that exhibits increased or decreased

XX PT expression in lung cancer

XX PS Claim 27; Page 288; 453pp; English.

XX CC The invention relates to a method for detecting a lung cancer-associated

XX CC transcript in a cell from a patient, comprising contacting a biological

XX CC sample from the patient with a polynucleotide that selectively hybridises

XX CC to a sequence that is at least 80 % identical to a gene that exhibits

XX CC increased or decreased expression in lung cancer samples. Lung

XX CC cancer-associated polynucleotides and polypeptides are used for

XX CC identifying a compound that modulates a lung cancer-associated

XX CC polypeptide, for inhibiting proliferation of a lung cancer-associated

XX CC cell to treat lung cancer in a patient and for treating a mammal having

XX CC lung cancer by administering a modulatory compound identified. The

Best Local Similarity 100.0%; Pred. No. 3.4e-169;

Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YASLAANAEFCNLFRMDNQGNGNVFSSLSLFAALALVRLGAQDLSLQDKLHV 60

DB 1 YASLAANAEFCNLFRMDNQGNGNVFSSLSLFAALALVRLGAQDLSLQDKLHV 60

QY 61 NTASGYGNSNSQSGLSQSKRVFSDINASHKDYLSIVNGLFAEKVYGFHKDYIECAEK 120

DB 61 NTASGYGNSNSQSGLSQSKRVFSDINASHKDYLSIVNGLFAEKVYGFHKDYIECAEK 120

QY 121 LYDAKVERVDFTNHLDTRENINKWVENETHGKIKNVIGEGGSISSAAMVLVNAVYFKG 180

DB 121 LYDAKVERVDFTNHLDTRENINKWVENETHGKIKNVIGEGGSISSAAMVLVNAVYFKG 180

QY 181 WQAPFTKSTTINCHFPKPSGKAVAMVHQERKFNLSVIEDPSMKILEYNGSINWYL 240

DB 181 WQAPFTKSTTINCHFPKPSGKAVAMVHQERKFNLSVIEDPSMKILEYNGSINWYL 240

QY 241 LPENDLSBIENKLTQNLMEWTNPRMTSKYVEVFPQFKIEKNYEMQYLRALGLKDF 300

DB 241 LPENDLSBIENKLTQNLMEWTNPRMTSKYVEVFPQFKIEKNYEMQYLRALGLKDF 300

QY 301 DESKADLSGIASGGRLYISRMHKSIIIVTEGTEATAATGTSNIVEKOLPQSTLFRADHP 360

DB 301 DESKADLSGIASGGRLYISRMHKSIIIVTEGTEATAATGTSNIVEKOLPQSTLFRADHP 360

QY 361 FLVIRKDDIILFSGKVSCP 380

DB 361 FLVIRKDDIILFSGKVSCP 380

RESULT 6

AAG64286

ID AAG64286 standard; Protein; 380 AA.

XX AC AAG64286;

XX DT 21-SEP-2001 (first entry)

XX DE Rat meginin protein.

XX KW Rat; meginin; renal mesangial cell; mesangium proliferative nephritis.

XX OS Rattus norvegicus.

XX PN WO200148019-A1.

XX PD 05-JUL-2001.

XX PF 26-DEC-2000; 2000WO-JP09251.

XX PR 28-DEC-1999; 99JP-0373677.

XX PA (KURO/) KUROKAWA K.

XX PA (MIYA/) MIYATA T.

XX PI Miyata T;

XX DR WPI; 2001-425651/45.

XX DR N-PSDB; AAH48181.

XX PT New antibody recognizing a partial sequence of rat meginin protein for

XX PT diagnosis of mesangium proliferative nephritis

XX PS Disclosure; Page 54-56; 63pp; Japanese.

XX CC The present invention relates to a novel antibody which recognises a

XX CC peptide consisting of residues 341-354 of rat meginin protein. The present

XX CC sequence is the protein sequence for rat meginin, which was used in the

XX CC present invention. Meginin is highly expressed in renal mesangial cells

XX CC and its level is elevated in mesangium proliferative nephritis. Assay of

XX CC the serum or urine level using the antibody is therefore indicative of

Query Match 100.0%; Score 1959; DB 24; Length 380;

CC this type of disorder.

XX Sequence 380 AA;

SQ Query Match 76.0%; Score 1488; DB 22; Length 380;

Best Local Similarity 73.7%; Pred. No. 2e-126;

Matches 280; Conservative 57; Mismatches 43; Indels 0; Gaps 0;

QY 1 MASLAANAEEFCNLFREMDNQNGNVFSSLSLFAALALVELGAQDLSQIDKLLHV 60
 DB 1 MASLAANAEEFCNLFREMDNQNGNVFSSLSLFAALALVELGAQDLSQIDKLLHV 60
 QY 61 NTASGYGNSNSQSGLOQLKRVFSDINASHKDYDLISVNGLPFAEKVYGFHKDYIECAEK 120
 DB 61 ISPSRQNGNSNSQSGLOQLKRVFSDINASHKDYDLISVNGLPFAEKVYGFHKDYIECAEK 120
 QY 121 LYDAKVERVDFTHLEDRNRINKWVENETHGKIKNVIKGGISSAVMVLVNAVYFKGK 180
 DB 121 LYNAKVERVDFTHLEDRNRINKWVENETHGKIKNVIKGGISSAVMVLVNAVYFKGK 180
 QY 181 WQSAFTKSETINCHFKSPKCGKAVAMHQRKFNLSVIEDPSMKILELYNGGINMYVL 240
 DB 181 WKSAPTKSDTSLCHFRSPSPGKAVNMHQRERFNLSTIQEPPMQILBLQYHGGISMYIM 240
 QY 241 LPENDLSEIENKLTQNLMWENPRMTSKYVEFFQPKIEKNYEMKQYLPALGLKDIF 300
 DB 241 LPEDDLSEIESKLSQNLMDWNSRKMKQYVNVFLPQPKIEKDYEMSHLSKVGLEDIF 300
 QY 301 DESKADLSGIASGGRLYVSKLMHKSLEIYSESGTEATAATGSNIVEKQLPQSTLFRADHP 360
 DB 301 VESRADLSGIASGGRLYVSKLMHKSLEIYSESGTEATAATGSNIVEKQLPQSTLFRADHP 360
 QY 361 FLFVIRKDDIILFSPKVSQCP 380
 DB 361 FLFVIRKNGIILFTGKVSCP 380

RESULT 7

AAV08255

ID AAY08255 standard; Protein; 380 AA.

XX AC AAY08255;

XX DT 14-JUL-1999 (first entry)

XX DE Rat megin protein.

XX KW Megsin; mesangial cell; treatment; diagnosis; disease; IgA nephropathy;

XX KW human; rat; murine.

XX OS Rattus rattus.

XX XX WO9915652-A1.

XX PD 01-APR-1999.

XX PF 22-SEP-1998; 98WO-JP04269.

XX PR 22-SEP-1997; 97UP-0275302.

XX XX (KURO/) KUROKAWA K.

XX PA (MIYA/) MIYATA T.

XX XX Miyata T;

XX DR WPI; 1999-276983/23.

XX DR N-P8DB; AAX56712.

XX PT Megsin protein expressed specifically in mesangial cells

XX PS Claim 1; Page 69-72; 100pp; Japanese.

XX CC This invention describes the isolation of novel megin nucleic acid and

CC proteins from human, rat and mouse tissue. This protein is expressed
 CC specifically in mesangial cells. The products of the invention are
 CC useful for the treatment and diagnosis of diseases involving mesangial
 CC cells, such as IgA nephropathy.

XX Sequence 380 AA;

SQ Query Match 75.4%; Score 1477; DB 20; Length 380;

Best Local Similarity 73.4%; Pred. No. 2e-125;

Matches 279; Conservative 56; Mismatches 45; Indels 0; Gaps 0;

QY 1 YASLAANAEEFCNLFREMDNQNGNVFSSLSLFAALALVELGAQDLSQIDKLLHV 60
 DB 1 MASLAANAEEFCNLFREMDNQNGNVFSSLSLFAALALVELGAQDLSQIDKLLHV 60
 QY 61 NTASGYGNSNSQSGLOQLKRVFSDINASHKDYDLISVNGLPFAEKVYGFHKDYIECAEK 120
 DB 61 ISPSRQNGNSNSQSGLOQLKRVFSDINASHKDYDLISVNGLPFAEKVYGFHKDYIECAEK 120
 QY 121 LYDAKVERVDFTHLEDRNRINKWVENETHGKIKNVIKGGISSAVMVLVNAVYFKGK 180
 DB 121 LYNAKVERVDFTHLEDRNRINKWVENETHGKIKNVIKGGISSAVMVLVNAVYFKGK 180
 QY 181 WQSAFTKSETINCHFKSPKCGKAVAMHQRKFNLSVIEDPSMKILELYNGGINMYVL 240
 DB 181 WKSAPTKSDTSLCHFRSPSPGKAVNMHQRERFNLSTIQEPPMQILBLQYHGGISMYIM 240
 QY 241 LPENDLSEIENKLTQNLMWENPRMTSKYVEFFQPKIEKNYEMKQYLPALGLKDIF 300
 DB 241 LPEDDLSEIESKLSQNLMDWNSRKMKQYVNVFLPQPKIEKDYEMSHLSKVGLEDIF 300
 QY 301 DESKADLSGIASGGRLYVSKLMHKSLEIYSESGTEATAATGSNIVEKQLPQSTLFRADHP 360
 DB 301 VESRADLSGIASGGRLYVSKLMHKSLEIYSESGTEATAATGSNIVEKQLPQSTLFRADHP 360
 QY 361 FLFVIRKDDIILFSPKVSQCP 380
 DB 361 FLFVIRKNGIILFTGKVSCP 380

RESULT 8

AAV24150

ID AAB24150 standard; Protein; 380 AA.

XX AC AAB24150;

XX DT 30-JAN-2001 (first entry)

XX DE Rat megin protein sequence SEQ ID NO:19.

XX KW Megsin; mesangium-predominant gene; serpin regulated; nephropathy;

XX KW IgA; immunoglobulin A; detection; renal function; renal disorder;

XX KW diagnosis; biological sample; blood; urine.

XX XX Rattus norvegicus.

XX FE Key Location/Qualifiers

XX FT Misc-difference 51

XX FT /note= "unspecified"

XX FT Misc-difference 94

XX FT /note= "unspecified"

XX XX WO20057189-A1.

XX XX 28-SEP-2000.

XX EF 17-MAR-2000; 2000WO-JP01646.

XX XX 19-MAR-1999; 99JP-0075305.

XX PR 28-OCT-1999; 99UP-0306623.

XX XX (KURO/) KUROKAWA K.

XX PA (FUSO) FUSO PHARM INC LTD.

```
PA (MIYA/) MIYATA T.
XX PI Miyata T;
XX PD WPI; 2000-611642/58.
XX DR N-PSDB; AAC55238.
XX PT Evaluating renal function comprises assaying megin protein in
XX PT biological sample -
XX PS Example 2; Page 81-84; 93pp; Japanese.
XX CC The present invention describes a method for evaluating renal function.
XX CC The method comprises assaying megin protein in biological sample. Also
XX CC described are: (1) use of a anti-megsin protein antibody for diagnosing
XX CC renal function; and (2) a kit for detecting megin protein comprising:
XX CC (a) anti-megsin protein antibody attached to solid magnetic particles;
XX CC (b) direct or indirect fixing for the antibody to the particles; and
XX CC (c) a magnet. The process is useful for evaluating renal function and
XX CC diagnosing renal disorders by assaying megin protein in biological
XX CC samples (preferably urine or blood). The process is reproducible and
XX CC gives accurate results. The present sequence represents the rat megin
XX CC protein, which is given in the exemplification of the present invention.
XX SQ Sequence 380 AA;

Query Match 75.4%; Score 1477; DB 21; Length 380;
Best Local Similarity 73.4%; Pred. No. 2e-125;
Matches 279; Conservative 56; Mismatches 45; Indels 0; Gaps 0;

QY 1 MASLAANAEEFCNLFREMDNNGNNGVFFSSLSLFAALALVRLGAQDDSLSQIDKLHY 60
DB 1 MASLAANAEEFGDFLREMDSSQNGNNGVFFSSLSLFTALSIRLGARGDCXQIDKALHF 60

QY 61 NTASGYGNSNSQSGLOQLKRVFSDINASHKDYDLSIVNGLFAEKVYGFHKDYIECAEK 120
DB 61 ISPSRQNSNSQGLGLOQLKRVLDINSHDKXKLSIANGVFAEKVDFHKSMECAEN 120

QY 121 LYDAKVERVDFTNHLDETRRNINKWVENETHGKIKNVIIGSGISSAVVWLVNAVYFGK 180
DB 121 LYNKAKVERVDFTNDETRFKINKNIENETHGKIKVLDGSSLSSSAVVWLVNAVYFGK 180

QY 181 WSAFTKSTIRCHPKSGKAVAMHQRKFNLSVIEDPSMKILEIRYNGINMYL 240
DB 181 WSAFTKSTIRCHPKSGKAVAMHQRKFNLSVIEDPSMKILEIRYNGINMYL 240

QY 241 LPENDLSEIKLTQNLMEWTNPRMTSKYVEVFFQPKIEKNYEMKQYLEALGLKDI 300
DB 241 LPEDDLSEIESKLSFQNLMDWTNSRKMSQYVNVFLPQPKIEKDYEMSHLSKVGLEDIF 300

RESULT 9
AAB83076
ID AAB83076 standard; Protein; 380 AA.
AC AAB83076;
XX 10-JUL-2001 (first entry)
XX Rat megin protein.
XX Rat; megin; mesangial cell proliferative nephritis; nephrotropic;
XX transgenic mouse; glomerular disease; animal model; drug screening.
XX Os Rattus norvegicus.
```

```
XX WO200124628-A1.
XX PD 12-APR-2001.
XX PF 05-OCT-2000; 2000WO-JP06988.
XX PR 05-OCT-1999; 99JP-0285736.
XX PA (KURO/) KUROKAWA K.
XX PA (MIYA/) MIYATA T.
XX PI Miyata T;
XX DR WPI; 2001-300136/31.
XX DR N-PSDB; AAF82439.
XX PT Mouse model for mesangial cell proliferative nephritis for development
XX PT and screening of new treatments -
XX PS Disclosure; Page 48-50; 62pp; Japanese.
XX CC The present sequence is rat megin. The human megin coding
XX CC sequence may be introduced into a mouse to produce an animal model of
XX CC mesangial cell proliferative nephritis. The symptoms include
XX CC enlargement of the mesangial base region, sedimentation of an immune
XX CC complex and an increase in mesangial cells. The animal model is useful
XX CC for analysing the pathology of chronic glomerular diseases and for
XX CC screening compositions for prevention and treatment of the diseases.
XX CC Highly uniform models can be made easily and in large numbers using
XX CC this method.
XX SQ Sequence 380 AA;

Query Match 75.4%; Score 1477; DB 22; Length 380;
Best Local Similarity 73.4%; Pred. No. 2e-125;
Matches 279; Conservative 56; Mismatches 45; Indels 0; Gaps 0;

QY 1 MASLAANAEEFCNLFREMDNNGNNGVFFSSLSLFAALALVRLGAQDDSLSQIDKLHY 60
DB 1 MASLAANAEEFGDFLREMDSSQNGNNGVFFSSLSLFTALSIRLGARGDCXQIDKALHF 60

QY 61 NTASGYGNSNSQSGLOQLKRVFSDINASHKDYDLSIVNGLFAEKVYGFHKDYIECAEK 120
DB 61 ISPSRQNSNSQGLGLOQLKRVLDINSHDKXKLSIANGVFAEKVDFHKSMECAEN 120

QY 121 LYDAKVERVDFTNHLDETRRNINKWVENETHGKIKNVIIGSGISSAVVWLVNAVYFGK 180
DB 121 LYNKAKVERVDFTNDETRFKINKNIENETHGKIKVLDGSSLSSSAVVWLVNAVYFGK 180

QY 181 WSAFTKSTIRCHPKSGKAVAMHQRKFNLSVIEDPSMKILEIRYNGINMYL 240
DB 181 WSAFTKSTIRCHPKSGKAVAMHQRKFNLSVIEDPSMKILEIRYNGINMYL 240

QY 241 LPENDLSEIKLTQNLMEWTNPRMTSKYVEVFFQPKIEKNYEMKQYLEALGLKDI 300
DB 241 LPEDDLSEIESKLSFQNLMDWTNSRKMSQYVNVFLPQPKIEKDYEMSHLSKVGLEDIF 300

QY 301 DESKADLSGIASGGLYISRMHKSIVIEVTEGTEATAATGSNIVEKQLPOSTLFRADHP 360
DB 301 VESRADLSGIASGGLYVSKLMHKSIIIEVSEEGTEATAATESNIVEKLLPSTVFRADRP 360

QY 361 FLFVIRKDDIILFSGKVSCEP 380
DB 361 FLFVIRKNGIILFTGKVSCEP 380

RESULT 10
AAY08256
ID AAY08256 standard; Protein; 368 AA.
AC AAY08256;
XX
```

DT 14-JUL-1999 (first entry)
 XX Mouse megin protein.
 DE Megin; mesangial cell; treatment; diagnosis; disease; IGA nephropathy;
 XX human; rat; murine.
 KW Mus musculus.
 XX
 OS Key Location/Qualifiers
 XX Protein 1.368
 FT /note= "partial sequence"
 XX
 XX W09515652-A1.
 XX
 XX 01-APR-1999.
 XX
 XX 22-SEP-1998; 98WO-JP04269.
 XX
 XX 22-SEP-1997; 97JP-0275302.
 XX
 XX (KURO/) KUROKAWA K.
 XX (MIYA/) MIYATA T.
 XX Miyata T;
 XX
 XX WPI; 1995-276983/23.
 XX N-PSDB; AAX56714.
 XX
 XX Megin protein expressed specifically in mesangial cells
 XX
 XX Claim 1; Page 76-79; 100pp; Japanese.
 XX
 XX This invention describes the isolation of novel megin nucleic acid and
 CC proteins from human, rat and mouse tissue. This protein is expressed
 CC specifically in mesangial cells. The products of the invention are
 CC useful for the treatment and diagnosis of diseases involving mesangial
 CC cells, such as IGA nephropathy.
 XX
 XX Sequence 368 AA;

Query Match 74.4%; Score 1458; DB 20; Length 368;
 Best Local Similarity 73.4%; Pred. No. 1e-123;
 Matches 270; Conservative 53; Mismatches 45; Indels 0; Gaps 0;
 QY 13 ENLFREMDNCGNVPFSSLSLPAALVRLGAQDDSLQSDKLLHNTASGYGNSNS 72
 Db 1 FDLFREMDSSQGNVPFSSLSLFTALTFLRLGARGECARQIDKALHFNIPRQGNSSNN 60
 QY 73 QGGLQSLKRVFSDINASHKDYDLSIVNGLFAEKVYGFHKDYIECAKLYDAKVERVDFT 132
 Db 61 QPGLQYQLKRVLADINSSHKDYELSIATGVFAEKVYDFHKYIECAENLVNAKVERVDFT 120
 QY 133 NHELTRNINIKWVENETHGKIKKVLGDSLSLSSA VVWLVNAVYFKGWSAFTKSETIN 192
 Db 121 NDVQDTRFKINKWVENETHGKIKKVLGDSLSLSSA VVWLVNAVYFKGWSAFTKSETIN 180
 QY 193 CFFKSPKSGKAVAMHQRKFNLSVIEDPSMKILRLYNGINGINMYVLLPENDLSEIENK 252
 Db 181 CFFRSFTCPGKVVMMHQRKFNLSVIEDPSMKILRLYNGINGINMYVLLPENDLSEIENK 240
 QY 253 LTFQNLMEWTNPRMTSKVEVFPQPKIEKNYEMQYLRALGLKDFBESKADLSGTAS 312
 Db 241 LSFQNLMDWTNRKMKSVQVFLDPQKLENTVMTHLKSLSGLKDFBESKADLSGTAS 300
 QY 313 GGRLVSRMKHSYEVTEGTEATAATGSNIVEKQLFQSTLFRADHPFLFVIRKDDIIL 372
 Db 301 GGRLVSKLMHKSFLVSWSEGEATAATGSNIVEKQLFQSTLFRADHPFLFVIRKDDIIL 360
 QY 373 FSGKVSCP 380
 Db 361 FTGKVSCP 368

RESULT 11
 AAB24151
 ID AAB24151 standard; Protein; 368 AA.
 XX
 AC AAB24151;
 XX
 ET 30-JAN-2001 (first entry)
 XX
 XX Mouse megin protein sequence SEQ ID NO:21.
 XX
 KW Megin; mesangium-predominant gene; serpin regulated; nephropathy;
 KW IGA; immunoglobulin A; detection; renal function; renal disorder;
 KW diagnosis; biological sample; blood; urine.
 XX
 OS Mus musculus.
 XX
 XX W0200057189-A1.
 XX
 XX 28-SEP-2000.
 XX
 XX 17-MAR-2000; 2000WO-JP01646.
 XX
 XX 19-MAR-1999; 99JP-0075305.
 XX 28-OCT-1999; 99JP-0306623.
 XX
 XX (KURO/) KUROKAWA K.
 XX (FUSO) FUSO PHARM IND LTD.
 XX (MIYA/) MIYATA T.
 XX Miyata T;
 XX
 XX WPI; 2000-611642/58.
 XX N-PSDB; AAC55239.
 XX
 XX Evaluating renal function comprises assaying megin protein in
 PT biological sample -
 XX
 XX Disclosure; Page 89-91; 93pp; Japanese.

XX The present invention describes a method for evaluating renal function.
 CC The method comprises assaying megin protein in biological sample. Also
 CC described are: (1) use of an anti-megin protein antibody for diagnosing
 CC renal function; and (2) a kit for detecting megin protein comprising:
 CC (a) anti-megin protein antibody attached to solid magnetic particles;
 CC (b) direct or indirect fixing for the antibody to the particles; and
 CC (c) a magnet. The process is useful for evaluating renal function and
 CC diagnosing renal disorders by assaying megin protein in biological
 CC samples (preferably urine or blood). The process is reproducible and
 CC gives accurate results. The present sequence represents the mouse megin
 CC protein, which is given in the exemplification of the present invention.
 XX
 XX Sequence 368 AA;
 Query Match 74.4%; Score 1458; DB 21; Length 368;
 Best Local Similarity 73.4%; Pred. No. 1e-123;
 Matches 270; Conservative 53; Mismatches 45; Indels 0; Gaps 0;
 QY 13 ENLFREMDNCGNVPFSSLSLPAALVRLGAQDDSLQSDKLLHNTASGYGNSNS 72
 Db 1 FDLFREMDSSQGNVPFSSLSLFTALTFLRLGARGECARQIDKALHFNIPRQGNSSNN 60
 QY 73 QGGLQSLKRVFSDINASHKDYDLSIVNGLFAEKVYGFHKDYIECAKLYDAKVERVDFT 132
 Db 61 QPGLQYQLKRVLADINSSHKDYELSIATGVFAEKVYDFHKYIECAENLVNAKVERVDFT 120
 QY 133 NHELTRNINIKWVENETHGKIKKVLGDSLSLSSA VVWLVNAVYFKGWSAFTKSETIN 192
 Db 121 NDVQDTRFKINKWVENETHGKIKKVLGDSLSLSSA VVWLVNAVYFKGWSAFTKSETIN 180
 QY 193 CFFKSPKSGKAVAMHQRKFNLSVIEDPSMKILRLYNGINGINMYVLLPENDLSEIENK 252
 Db 181 CFFRSFTCPGKVVMMHQRKFNLSVIEDPSMKILRLYNGINGINMYVLLPENDLSEIENK 240

QY	133	NHLEDTRNINKWENETHGKIKNVLGEGIGSSAVMVLINAVYFKGKWSAFTKSETIN	193
Db	121	NDVQDTAFKLNKWIENETHGKIKKVLGDSSSSSAVMVLINAVYFKGKWSAFTKTDTJS	180
QY	193	CHFKSPKSGKAVAMHGERKFNLSVIEDPSSMKILELRVNGGINMVLPPNDLSBIENK	252
Db	181	CFRSPFCFKGVNMWHQSERFNLSTIQPPNVOVLEQYHGGISWVIMLPDGLCHIESK	240
QY	253	LTFQNLMEWNPFRMTSKYEVFFPQFKIERKNYEMKQYLAALGLKDIIDFESKADLSGIAS	312
Db	241	LGFQNLMDWNRKMKSGYVNVLPQFKIEKNYEMTHLKSGLKDIIDFESSADLSGIAS	300
QY	313	GRRLYSMMHMKSYIEVTEBGTENTATAGSNIVEKQLPPOSTLFRADHFLFVIRKDDIIL	372
Db	301	GRRLYSVKIMKHSFIEVSEBGTENTATENNIVEKQLPPESTVFRADRFLEFVIRKDDIIL	360
QY	373	FGSKVSCP 380	
Db	361	FGTKVSCP 368	
RESULT 13			
ABG76511			
ID	ABG76511	standard; Protein; 378 AA.	
AC	ABG76511;		
XX	05-NOV-2002	(first entry)	
XX		DNA encoding protein modification and maintenance molecule #15.	
DE		Protein modification and maintenance molecule; gastrointestinal disorder;	
XX		cyphagata; esophageal spasm; gastritis; anorexia; nausea; hypertension;	
KW		cardiovascular disorder; atherosclerosis; vasculitis; aneurysm; allergy;	
KW		ischemic heart disease; autoimmune disorder; inflammatory disorder;	
KW		acquired immunodeficiency syndrome; AIDS; ankylosing spondylitis; cancer;	
KW		anemia; amyloidosis; cell proliferative; arteriosclerotic bursitis;	
KW		cirrhosis; developmental disorder; renal tubular acidosis; anaemia;	
KW		bone resorption; epilepsy; epithelial disorder; keratosis pilaris;	
KW		allergic contact dermatitis; insect bite; keloid; dermatofibroma; eczema;	
KW		neurological disorder; stroke; cerebral neoplasm; Alzheimer's disease;	
KW		Huntington's disease; dementia; reproductive disorder; infertility;	
KW		endometriosis; gynaecostasia; ectopic pregnancy; gene therapy.	
XX		Homo sapiens.	
OS			
XX		WO200260942-A2.	
PN		08-AUG-2002.	
PD			
XX		30-JAN-2002; 2002WO-USG2813.	
PF			
XX		31-JAN-2001; 2001US-265705P.	
PR		05-FEB-2001; 2001US-266762P.	
PR		16-FEB-2001; 2001US-269581P.	
PR		23-FEB-2001; 2001US-271198P.	
PR		01-MAR-2001; 2001US-272813P.	
PR		13-MAR-2001; 2001US-275586P.	
PR		23-MAR-2001; 2001US-278505P.	
PR		30-MAR-2001; 2001US-280539P.	
XX		(INCYTE GENOMICS INC.	
PA			
XX		Warren BA, Honchell CD, Lu Y, Wallia NK, Burford N, Delegeane AM;	
PI		Gandhi AS, Baugan MR, Griffin JA, Gietzen KJ, Lu DAM, Ison CH;	
PI		Ramkumar J, Tang TY, Ial PG, Borowski ML, Duggan BM, Hafalia AUA;	
PI		Arvizu C, Thangavelu K, Yao MG, Elliott VS, Ding L, Yue H, Lee S;	
PI		Swarnakar A, Tran UK, Xu Y;	
DR		WPI; 2002-608499/65.	
DR		N-FSDB; AB958382.	
XX			

QY	114	YIECAKYLDAKVERVDFTNHLEDTNRINIKWVETETGKIKNVI GEGISSAVMLVN	173
DB	121	YUSCEKRYQARLQTVDFEQSTEETRKMIKAWYENKTKGVANLFGKSTIDFSSVYLVN	180
QY	174	AVYFKGQWQSAFTKSETINCFXGPKCSGK--AVAMEQERKFNLSVIEDSMKILELY	231
DB	181	TIYERQQRNKP-----QGKNTVEMVYQIGTFKLAIVKEPQMOVLELPY	225
QY	232	-NGGINMVILPEN--DLSEINKUTTCNLMWTFNPRMTSKYVYFFQFKIKKNYEMK	288
DB	226	VNNKLSMILLPVGIVANLKQIEKQKNSGTSFHEWTSSNNMEREVEVHLPRFLRIKYELN	285
QY	289	QYLRALGLKLDIPDSSKADLSGASGRIYLSRMHKSIVETEBGTETAATGSMIVKEQ	348
DB	286	SLLKPLGVYDIPNOVQADLSGWSPTKGLIYSKAIHKSILDVSEEGTEAAATGDSIAVKS	345
QY	349	LPOSTLFRADHPFLFVIR--KODIILFSGKVSCP	380
DB	346	LPMQAQKANHPFLFFIRHTEHTNTILFCGKLASP	379
RESULT 15			
ID	AAW47207		
AC	AAW47207 standard; Protein; 395 AA.		
AC	AAW47207;		
DT	12-FEB-2002 (first entry)		
XX	Human NOV1 protein.		
DE	Human; NOV1; metabolic disorder; neurodegenerative disorder;		
XX	immune disorder; haematoepoietic disorder; developmental disease; cancer;		
KW	retinal disease; feeding disorder; vaccine; infection; gene therapy;		
KW	neurological disorder; psychotic disorder; G-protein coupled receptor;		
KW	cytostatic; antidiabetic; viricide; neuroprotective; nootropic;		
KW	analgesic; antidepressant; antimigraine; anticonvulsant; neuroleptic;		
KW	antiasthmatic; antiallergic; antiinflammatory; anorectic; antiarthritic;		
KW	antipsychotic; antitachycardic; antibacterial; fungicide;		
KW	osteopathic; prozooscide; antulcer; hypertensive; hypotensive;		
KW	antinfertility; vulnerary; nephrotropic; antilipemic; leupin;		
XX	chromosome 18.		
OS	Homo sapiens.		
XX	WC200174851-A2.		
PN	11-OCT-2001.		
PD	30-MAR-2001; 2001WO-US10039.		
PF	30-MAR-2000; 2000US-193205P.		
XX	30-MAR-2000; 2000US-193339P.		
PR	05-APR-2000; 2000US-195343P.		
PR	06-APR-2000; 2000US-195005P.		
PR	06-APR-2000; 2000US-195083P.		
PR	10-APR-2000; 2000US-195792P.		
PR	11-APR-2000; 2000US-196556P.		
PR	13-APR-2000; 2000US-197081P.		
PR	14-APR-2000; 2000US-197087P.		
PR	14-APR-2000; 2000US-197525P.		
PR	29-MAR-2001; 2001US-0823187.		
XX	(CURA-) CURAGEN CORP.		
DA	Majumder X, Spaderna SK, Taupier RU, Padigaru M, Burgess CE;		
XX	Shimkets RA, Spytek KA, Liu X, Patturajan M, Gusev VV;		
PI	WPI: 2001-626379/72.		
DR	N-PSDB; A3A01980.		
DR	New G protein-coupled receptor related polypeptides and polynucleotides		
PT	for diagnosis, prevention and treatment of metabolic,		

PT neurodegenerative, retinal, immune, hematopoietic disorders, diabetes,
 XX obesity and infections -
 XX
 XX
 XX Claim 1; Page 9; 194pp; English.
 CC The present invention provides the protein and coding sequences pf novel
 CC human G-protein coupled receptors, designated NOV1, NOV2, NOV3, NOV4,
 CC NOV5a, NOV5b, NOV5a, NOV6b, NOV7, NOV8a, NOV8b, NOV9 and NOV10. These can
 CC be used in the treatment of NOVX related diseases, including cancer,
 CC metabolic, neurodegenerative, immune, hematopoietic, developmental,
 CC retinal, feeding, neurological and psychotic diseases and disorders and
 CC infections. The present sequence is the NOV1 protein, the gene for which
 CC is found on chromosome 18. The NOV1 protein shares homology with the
 CC leupin protein.
 XX
 XX Sequence 395 AA;
 SQ
 Query Match 39.4%; Score 771.5; DB 22; Length 395;
 Best Local Similarity 40.3%; Pred. No. 2.6e-61;
 Matches 159; Conservative 89; Mismatches 132; Indels 15; Gaps 5;
 QY 1 MASLAAANAEFCNLPREDDNQNGVFFSLSLFAALVRLGAQDLSIDKLHLV 60
 Db 1 MDSLVTANTKFCDFLQETIGKDDKHKNFFSPISLSAALGMVRLGARDSAHQIDVLFH 60
 QY 61 NTASGY-----GNSSNSAGSLQS-QLKRVFSDINASHKDYDLSIVNGLPAAKYVGKHY 114
 Db 61 NBFQSNESKPPAGLNSNSGLVSCYFQQLSKLDRINTYTLISIANRUYGEQFFPCQBY 120
 QY 115 IECAKKLVDAAKRVDFTNHLEDTRNNKNVNEETHGKIKVIGEGIGSSAAWVLVNA 174
 Db 121 LGVTFPHITISVDFQKNPEKSRQEIFNPWECOSGKIKOLFSDAINAEIVLVNA 180
 QY 175 VYFAGKWSAFTKSETINGHFKSPKSGKAVAMHQBRKPNLSVIEDPSMKLLELY-NG 233
 Db 181 VYFKAKWTFVDHENTVDAPFCLNQENESKVRMTQGLYRIGFIEYKAGLLEMYTNG 240
 QY 234 GINMYVILPEND-----LSLIEKLTFQNLNWTNPRMTSKYVEVFPQKLEKNVEM 287
 Db 241 KLSNFTVLPSSKXNKLGLSELEKTYIEKKVAMSSSENNSVLTSPFRFTLESDYDL 300
 QY 288 KQVRLALGLKIDFESKADLGGIASGGRILYSRMMHKSYTEVEEGTATAAGTSGNIVEK 347
 Db 301 NSLIQDMGIIITDFETRADLTGTSPEPNILYLSKLIHKTFVEVDENGTOAAAGTAVVSR 360
 QY 348 QLFPQSTLPRADHPFLVIR--KDDIILFGKVCSP 380
 Db 361 SURSWYENANHPLEFFTRHNKTTOTILFVGRVCSP 395

New G protein-coupled receptor related polypeptides and polynucleotides for diagnosis, prevention and treatment of metabolic,

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 16:16:51 ; Search time 21 Seconds
(without alignments)
765.625 Million cell updates/sec

Title: US-09-508-997A-2
Perfect score: 1959
Sequence: 1 MASLAANAEEFCNLFREMD.....FLFWIRKDDIILFGKVSQP 380

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 segs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/aaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/aaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/aaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/aaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/aaa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/aaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1959	100.0	380	2	US-08-472-659-34
2	1959	100.0	380	2	US-08-474-661-34
3	1959	100.0	380	2	US-08-611-977-34
4	754.5	38.5	425	4	US-10-024-427-2
5	752	38.4	390	1	US-08-568-147B-2
6	747	38.1	390	4	US-09-266-910-3
7	745	38.0	390	4	US-09-266-910-4
8	730.5	37.3	375	1	US-08-121-714-8
9	730.5	37.3	375	1	US-08-477-108A-8
10	730.5	37.3	375	5	PCT-US93-08322-8
11	730.5	37.3	375	5	US-08-477-112-8
12	722.5	36.9	379	1	US-08-121-714-4
13	722.5	36.9	379	2	US-08-477-108A-4
14	722.5	36.9	379	2	US-08-477-112-4
15	722.5	36.9	379	5	PCT-US93-08322-4
16	715.5	36.5	391	4	US-09-123-912-110
17	715.5	36.5	391	4	US-09-643-597-110
18	715.5	36.5	391	4	US-09-480-884A-110
19	715.5	36.5	391	4	US-09-542-615A-110
20	715.5	36.5	391	4	US-09-606-421B-110
21	715	36.5	400	4	US-09-123-912-112
22	715	36.5	400	4	US-09-643-597-112
23	715	36.5	400	4	US-09-480-884A-112
24	715	36.5	400	4	US-09-542-615A-112
25	715	36.5	400	4	US-09-606-421B-112
26	673	34.4	382	1	US-07-768-286B-6
27	673	34.4	382	1	US-08-487-623B-3

28	673	34.4	382	2	US-08-997-040-3	Sequence 3, Appli
29	673	34.4	382	2	US-09-203-237-3	Sequence 3, Appli
30	671	34.3	392	1	US-07-768-286B-4	Sequence 4, Appli
31	664.5	33.9	415	1	US-07-911-531-19	Sequence 19, Appli
32	664.5	33.9	415	1	US-07-693-636A-19	Sequence 19, Appli
33	664.5	33.9	415	1	US-07-768-286B-2	Sequence 2, Appli
34	664.5	33.9	434	1	US-07-679-052A-15	Sequence 15, Appli
35	662.5	33.8	415	4	US-09-026-408-4	Sequence 2, Appli
36	662	33.8	374	1	US-08-464-148-2	Sequence 2, Appli
37	662	33.8	374	1	US-08-385-500-2	Sequence 2, Appli
38	662	33.8	374	1	US-08-846-784-2	Sequence 2, Appli
39	660.5	33.7	438	1	US-07-679-052A-17	Sequence 17, Appli
40	655	33.4	386	4	US-08-545-573A-39	Sequence 39, Appli
41	653.5	33.4	405	1	US-08-121-714-7	Sequence 7, Appli
42	653.5	33.4	405	1	US-08-477-108A-7	Sequence 7, Appli
43	653.5	33.4	405	2	US-08-477-112-7	Sequence 7, Appli
44	653.5	33.4	405	5	PCT-US93-08322-7	Sequence 7, Appli
45	652	33.3	409	4	US-09-613-303-55	Sequence 55, Appli

ALIGNMENTS

RESULT 1
US-08-472-659-34
; Sequence 34, Application US/08472659
; Patent No. 5831030
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Fuyuki
; APPLICANT: TSUJIMOTO, No. 583103000
; APPLICANT: NAKAZATO, Hiroshi
; APPLICANT: MIURA, Kenju
; APPLICANT: KURIHARA, Tatsuya
; APPLICANT: YAMAICHI, Kozo
; APPLICANT: YAMAGUCHI, No. 583103000
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,659
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5-067339
; FILING DATE: 04-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-248
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:

LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-472-659-34

Query Match 100.0%; Score 1959; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.3e-178;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASLAANAABCFNLFFEMDDNQNGNVFFSSLSLFAALALVRLGAQDDSLSDIKLLHV 60
DB 1 MASLAANAABCFNLFFEMDDNQNGNVFFSSLSLFAALALVRLGAQDDSLSDIKLLHV 60
QY 61 NTASGYGNSNSQSGLOSLKRVFSDINASHKDYDLSI VNGLFAEKVYGFHKDYIECAEK 120
DB 61 NTASGYGNSNSQSGLOSLKRVFSDINASHKDYDLSI VNGLFAEKVYGFHKDYIECAEK 120
QY 121 LYDAKVERVDFTNLEDTRENINKWVENETHGKIKNVI GEGGSSSAWVLYNAVYFKG 180
DB 121 LYDAKVERVDFTNLEDTRENINKWVENETHGKIKNVI GEGGSSSAWVLYNAVYFKG 180
QY 181 WQSAFTKSETINCHFKSPKCGKAVAMVHQRKFNLSVIEDPSMKILELRNGINMYVL 240
DB 181 WQSAFTKSETINCHFKSPKCGKAVAMVHQRKFNLSVIEDPSMKILELRNGINMYVL 240
QY 241 LPENDLSEIENKLTFOQLMEWTNPRMTSKYVEVFPQPKIEKNYEMQYLRALGLKIDIF 300
DB 241 LPENDLSEIENKLTFOQLMEWTNPRMTSKYVEVFPQPKIEKNYEMQYLRALGLKIDIF 300
QY 301 DESKADLSGIASGGRLYISRMHKSIVIEVTEEGTEATAATGNSIVEXKOLPQSTLFRADHP 360
DB 301 DESKADLSGIASGGRLYISRMHKSIVIEVTEEGTEATAATGNSIVEXKOLPQSTLFRADHP 360
QY 361 FLFVIRKDDIILFSGKVSCP 380
DB 361 FLFVIRKDDIILFSGKVSCP 380

RESULT 2

US-08-474-661-34
Sequence 34, Application US/08474661
Patent No. 5874253
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURUOKA, No. 5874253uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5874253uhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. 5874253om.
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,661
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028

FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: REA, TERESA STANEK
REGISTRATION NUMBER: 30,427
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-474-661-34

Query Match 100.0%; Score 1959; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.3e-178;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASLAANAABCFNLFFEMDDNQNGNVFFSSLSLFAALALVRLGAQDDSLSDIKLLHV 60
DB 1 MASLAANAABCFNLFFEMDDNQNGNVFFSSLSLFAALALVRLGAQDDSLSDIKLLHV 60
QY 61 NTASGYGNSNSQSGLOSLKRVFSDINASHKDYDLSI VNGLFAEKVYGFHKDYIECAEK 120
DB 61 NTASGYGNSNSQSGLOSLKRVFSDINASHKDYDLSI VNGLFAEKVYGFHKDYIECAEK 120
QY 121 LYDAKVERVDFTNLEDTRENINKWVENETHGKIKNVI GEGGSSSAWVLYNAVYFKG 180
DB 121 LYDAKVERVDFTNLEDTRENINKWVENETHGKIKNVI GEGGSSSAWVLYNAVYFKG 180
QY 181 WQSAFTKSETINCHFKSPKCGKAVAMVHQRKFNLSVIEDPSMKILELRNGINMYVL 240
DB 181 WQSAFTKSETINCHFKSPKCGKAVAMVHQRKFNLSVIEDPSMKILELRNGINMYVL 240
QY 241 LPENDLSEIENKLTFOQLMEWTNPRMTSKYVEVFPQPKIEKNYEMQYLRALGLKIDIF 300
DB 241 LPENDLSEIENKLTFOQLMEWTNPRMTSKYVEVFPQPKIEKNYEMQYLRALGLKIDIF 300
QY 301 DESKADLSGIASGGRLYISRMHKSIVIEVTEEGTEATAATGNSIVEXKOLPQSTLFRADHP 360
DB 301 DESKADLSGIASGGRLYISRMHKSIVIEVTEEGTEATAATGNSIVEXKOLPQSTLFRADHP 360
QY 361 FLFVIRKDDIILFSGKVSCP 380
DB 361 FLFVIRKDDIILFSGKVSCP 380

RESULT 3

US-08-611-977-34
Sequence 34, Application US/08611977
Patent No. 5972886
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURUOKA, No. 5972886uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. 5972886cmi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404

```

; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22133-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611,977
; FILING DATE: 06-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-067339
; FILING DATE: 04-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-204
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-611-977-34

Query Match 100.0%; Score 1959; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.3e-178;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASLAAANAECFNLPREMDNDNGNGNVFSSLSLPAALALVRLGAODDSLSQIDKLHV 60
DB 1 MASLAAANAECFNLPREMDNDNGNGNVFSSLSLPAALALVRLGAODDSLSQIDKLHV 60
QY 61 NTASGYGNSNSQSGLSQSKRVFSDINASHKDYDLSIVNGLFAEKYGFHDKYIECAEK 120
DB 61 NTASGYGNSNSQSGLSQSKRVFSDINASHKDYDLSIVNGLFAEKYGFHDKYIECAEK 120
QY 121 LYDAKVERVDFTHLEDTNRNINKNWVENETHGKIKNVIGEGGISSAVMLVNAVYFKG 180
DB 121 LYDAKVERVDFTHLEDTNRNINKNWVENETHGKIKNVIGEGGISSAVMLVNAVYFKG 180
QY 181 WSAFTKSETINCHFKSPKCSGKAVAMHQRKFNLSVIEDPSPMKILELRVNGGINVYL 240
DB 181 WSAFTKSETINCHFKSPKCSGKAVAMHQRKFNLSVIEDPSPMKILELRVNGGINVYL 240
QY 241 LPENDLSIENKLTQNLMEWTNPRMTSKYVEVFPQFKIKENYEMKQYLALGLKIDF 300
DB 241 LPENDLSIENKLTQNLMEWTNPRMTSKYVEVFPQFKIKENYEMKQYLALGLKIDF 300
QY 301 DRSKADLSGIASGRILYISRMHKSYLEVTEEGTEATAATGNSNIVEKQLPSTLFRADHP 360
DB 301 DESKADLSGIASGRILYISRMHKSYLEVTEEGTEATAATGNSNIVEKQLPSTLFRADHP 360
QY 361 FLFVIRKDDIILFSGKVSCP 380
DB 361 FLFVIRKDDIILFSGKVSCP 380

RESULT 4
US-10-024-427-2
; Sequence 2, Application US/1002427

; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22133-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611,977
; FILING DATE: 06-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-067339
; FILING DATE: 04-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-204
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-611-977-34

; Patent No. 6583269
; GENERAL INFORMATION:
; APPLICANT: Hu, Yi
; APPLICANT: Nepomnichy, Boris
; APPLICANT: Waite, D. Wade
; TITLE OF INVENTION: No. 6583269el Human Protease Inhibitor and Polynucleotides Enc
; FILE REFERENCE: LEX-0283-USA
; CURRENT APPLICATION NUMBER: US/10/024,427
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: US 60/256,287
; PRIOR FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 425
; TYPE: PRT
; ORGANISM: homo sapiens
; US-10-024-427-2

Query Match 38.5%; Score 754.5; DB 4; Length 425;
Best Local Similarity 37.4%; Pred. No. 1.4e-63;
Matches 159; Conservative 87; Mismatches 134; Indels 45; Gaps 5;

QY 1 MASLAAANAECFNLPREMDNDNGNGNVFSSLSLPAALALVRLGAODDSLSQIDKLHV 60
DB 1 MDSLVTANTKFCFDLFOEIGKDRHKNIFTSLSLSAALGMVRLGARSASAHQIDEVLHF 60
QY 61 NTAS-----GYGNSNSQSGLSQSKRVFSDINASHKDYDLSIVNGLFAEKYGFHDKYIECAEK 120
DB 61 NEFSQNSKBPDPCLKSNKOKVLADSSLEGQKKTTELPQQAGSLNNSGLVSCYFGQLL 120
QY 85 SDINASHKDYDLSIVNGLFAEKYGFHDKYIECAEKLYDAKVERVDFTHLEDTNRNIN 144
DB 121 SKLDRIKTDYTLISIANELYGEEFPIQCEYLDGVIQFYHTTIESVDFOKNPEKSRQEIF 180
QY 145 WVENETHGKIKNVIGEGGISSAVMLVNAVYFKGQWQAFKTSINCHFKSPKCSGKA 204
DB 181 WVCQSQGKIKELFSKDAINAEATVLVNAVYFKAKWETYPDENTVDAPFCLNANENKS 240
QY 205 VAMHQRKFNLSVIEDPSPMKILELRVNGGINVYLLPEND-----LSIENKLTQNL 257
DB 241 VKMTQKGYIRIGIEVRAQILEMRYTKGLSNVLLPSSHKNLKLGLBELEKRTIYER 300
QY 258 LMEWTNPRMTSKYVEVFPQFKIKENYEMKQYLALGLKIDFDESKADLSGIASGRILY 317
DB 301 MWAWSSNNSESVLSRPTLSDSYDANSILQDWGITDIFDETRADLTGSPSPNLY 360
QY 318 ISRMHKSYLEVTEEGTEATAATGNSNIVEKQLPSTLFRADHPFLFVIR--KDDIILFSG 375
DB 361 LSKTIHKTFVSDENGTOAAATGAWSERSLRSWVEFNANHPFLFTRHNTQTILPYG 420
QY 376 KVSCP 380
DB 421 RVSCP 425

RESULT 5
US-08-568-147B-2
; Sequence 2, Application US/08568147B
; Patent No. 5763422
; GENERAL INFORMATION:
; APPLICANT: Suminami, Yoshinori
; APPLICANT: Kato, Hiroshi
; APPLICANT: Sekiguchi, Kiyoshi
; APPLICANT: Takeda, Katsumichi
; TITLE OF INVENTION: DNA FRAGMENT CODING FOR SQUAMOUS CELL
; TITLE OF INVENTION: CARCINOMA-ASSOCIATED ANTIGEN
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York

```

/ COUNTRY: USA
/ ZIP: 11530
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/568,147B
/ FILING DATE:
/ CLASSIFICATION: 536
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 800,952
/ FILING DATE: 02-DEC-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Digiglio, Frank S.
/ REGISTRATION NUMBER: 31,346
/ REFERENCE/DOCKET NUMBER: 8425
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 516-742-4343
/ TELEFAX: 516-742-4366
/ TELEX: 230 901 SANS UR
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 390 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-568-147B-2

Query Match 38.4%; Score 752; DB 1; Length 390;
Best Local Similarity 41.5%; Pred. No. 2e-63;
Matches 145; Conservative 81; Mismatches 126; Indels 26; Gaps 10;

QY 1 MASLAANAECFNLPREMDNQNGNVFFSSLSFAALALVRLGAODSLSQIDKLIHV 60
DB 1 MNSLSEANTKMFDFLQOQPKSKEN-NIFYSPISITISALGMYLLGAKNTAQIKKVLHF 59

QY 61 -----NTASGYGNSNSQSG-LQSOLKRVFSDINASHKDYDLSIVNGLFAKVGFGHKDY 114
DB 60 DQVTNTTGAATYHVRSGNVHGFQKLLTEFNKSTDAVELKIANKLFGKTYLFLQY 119

QY 115 IECABKLDYAKVERVPTNHLNINWENETHGKIKNVIGEGGSSAVVAVNA 174
DB 120 LDAIKKFTQTSVESVDFANAPESPKKINSWVSQTNKIKNLIPEGNISNTTLVNA 179

QY 175 VYFKGWSAFTKSETINCHFKSPKCSKAVAMMHQERKNLSVIEDPSMKILELRNG- 233
DB 180 IYFKQWKKKEDTKKEKFWPNKNTYKSIQMRQYTSFHPASLEDVQAKVLEPYKKG 239

QY 234 GINMYVLLPE--NDLSEIENKLTQNLMEWTPRMTSKYVEVFFQFKIEKNYEMKOYL 291
DB 240 DUSNVLLPNEIDGLQKLEKTAELKMEWTSQNNRETRVDLHLPFKVSESYDLKDTL 299

QY 292 RALGKLDIPDSKADLSGIASGRGLYISRMHKSYYIEVTEGTATAT-----GSNIVE 346
DB 300 RWGMVDIFN-GDADLSGMTSGRGLVSGVLHAFVETVEGAAFAAATAVVGFGSS--- 355

QY 347 KQPOST--LFRADHPFLVIR--KDDIILFSGKVSCP 380
DB 356 ---PASTNEEHCHHPFLFFIQNKNTNSILFYGRFSSP 390

RESULT 6
US-08-568-147B-2
/ Sequence 3, Application US/09266910
/ Patent No. 6344362
/ GENERAL INFORMATION:
/ APPLICANT:
/ TITLE OF INVENTION: Use of a recombinant protein as receptor of a
/ TITLE OF INVENTION: hepatitis virus
/ NUMBER OF SEQUENCES: 8
/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/266,910
/ FILING DATE:
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 390 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ US-09-266-910-3

Query Match 38.1%; Score 747; DB 4; Length 390;
Best Local Similarity 41.2%; Pred. No. 6.1e-63;
Matches 164; Conservative 81; Mismatches 127; Indels 26; Gaps 10;

QY 1 MASLAANAECFNLPREMDNQNGNVFFSSLSFAALALVRLGAODSLSQIDKLIHV 60
DB 1 MNSLSEANTKMFDFLQOQPKSKEN-NIFYSPISITISALGMYLLGAKNTAQIKKVLHF 59

QY 61 -----NTASGYGNSNSQSG-LQSOLKRVFSDINASHKDYDLSIVNGLFAKVGFGHKDY 114
DB 60 DQVTNTTGAATYHVRSGNVHGFQKLLTEFNKSTDAVELKIANKLFGKTYLFLQY 119

QY 115 IECABKLDYAKVERVPTNHLNINWENETHGKIKNVIGEGGSSAVVAVNA 174
DB 120 LDAIKKFTQTSVESVDFANAPESPKKINSWVSQTNKIKNLIPEGNISNTTLVNA 179

QY 175 VYFKGWSAFTKSETINCHFKSPKCSKAVAMMHQERKNLSVIEDPSMKILELRNG- 233
DB 180 IYFKQWKKKEDTKKEKFWPNKNTYKSIQMRQYTSFHPASLEDVQAKVLEPYKKG 239

QY 234 GINMYVLLPE--NDLSEIENKLTQNLMEWTPRMTSKYVEVFFQFKIEKNYEMKOYL 291
DB 240 DLSMTVLLPNEIDGLQKLEKTAELKMEWTSQNNRETRVDLHLPFKVSESYDLKDTL 299

QY 292 RALGKLDIPDSKADLSGIASGRGLYISRMHKSYYIEVTEGTATAT-----GSNIVE 346
DB 300 RWGMVDIFN-GDADLSGMTSGRGLVSGVLHAFVETVEGAAFAAATAVVGFGSS--- 355

QY 347 KQPOST--LFRADHPFLVIR--KDDIILFSGKVSCP 380
DB 356 ---PASTNEEHCHHPFLFFIQNKNTNSILFYGRFSSP 390

RESULT 7
US-09-266-910-4
/ Sequence 4, Application US/09266910
/ Patent No. 6344362
/ GENERAL INFORMATION:
/ APPLICANT:
/ TITLE OF INVENTION: Use of a recombinant protein as receptor of a
/ TITLE OF INVENTION: hepatitis virus
/ NUMBER OF SEQUENCES: 8
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/266,910
/ FILING DATE:
/ INFORMATION FOR SEQ ID NO: 4:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 390 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single

```

;
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
;
US-09-266-910-4

Query Match      38.0%; Score 745; DB 4; Length 390;
Best Local Similarity 41.2%; Pred. No. 9.5e-63;
Matches 164; Conservative 81; Mismatches 127; Indels 26; Gaps 10;

QY 1 MASLAANAEPFNLFREMDNQGNGVFFSSLSLFAALALVRLGAQDSSLSQIDKLLHV 60
Db 1 MNSLSANTKFFDFQOQPRKSEN-NIFYSPISIIISALGMVLLGAKNTAQIKKVLHF 59

QY 61 -----NTASGYGNSNSQSG-LOSQKRVFSDINASHKDYDLSIVNGLPARKYVGHKDY 114
Db 60 DQVTEHTGKANTHVRSGNVHQQFKLTFEKNSTDAVELKIANKLFGEKTYLFLQSY 119

QY 115 IBOAEKLYDAKVERDFTNHLNEDTRNINKWVENETHGKIKNVIGEGISSAVMVLVNA 174
Db 120 LDAIKKFYQTSVESVDFANAPESPCKINSWVESQTNKIKXLIPEGNISNTVLVNA 179

QY 175 VYKKGWQSAFTKETINCHFKSPKCSKAVAMHQBKRNLSVIEDPSMKILELRNG- 233
Db 180 IYFKGWQEKFNKSDTKEEFKFNKNTYKSIQMRQVTSFHPASLEDVQAKVLEIDYK 239

QY 234 GINMYVLLPE--NLSIEINKLTFQNLMEWTNPRMTSKYVEVFPQPKIEKNYEMKQVL 291
Db 240 DLSMIVLLENEIDGLKLEKLTAKLMWETSLQWRETRVDLHLPRFKVEESYDLKOTL 299

QY 292 RAUGLKDIPDSKADLSGIASGRRLYISRMHKSXYIEVTEGTATAT-----GSNIVE 346
Db 300 RTMGWVDIFN-GDADLSGMTSGRLVSGVLHKAFFVEVIEEGAAATATAVAFSS--- 355

QY 347 KQIPOST--LFRADHPPLFVIR--KDDILFSGKVSPP 380
Db 356 ---PSTNEEHPCHNPPLEFFIRQNTKINSILFYGRFSPP 390

RESULT 8
US-08-121-714-8
; Sequence 8, Application US/08121714
; Patent No. 5470970
;
; GENERAL INFORMATION:
; APPLICANT: Sager, Ruth
; TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH
; TITILE OF INVENTION: TUMOR SUPPRESSING ACTIVITY
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/121,714
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/938,823
; FILING DATE: 09/01/92
; APPLICATION NUMBER: 07/844,296
; FILING DATE: 02/28/92
; APPLICATION NUMBER: 07/662,216
; FILING DATE: 02/28/91
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819

```

```

;
; REFERENCE/DOCKET NUMBER: 00530/072001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 375
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
;
US-08-121-714-8

Query Match      37.3%; Score 730.5; DB 1; Length 375;
Best Local Similarity 37.9%; Pred. No. 2.2e-61;
Matches 146; Conservative 85; Mismatches 135; Indels 19; Gaps 4;

QY 1 MASLAANAEPFNLFREMDNQGNGVFFSSLSLFAALALVRLGAQDSSLSQIDKLLHV 60
Db 1 MEQLSTANTHFADVLFRALNESDFTGNIFISPLSISSALAMIFLGTGNTAAQVSKALYP 60

QY 61 NTASGYGNSNSQSGLOSQKRVFSDINASHKDYDLSIVNGLPARKYVGHKDYIECAEK 120
Db 61 DTV-----EDHSRQSLNADINKPGAPILKANRLTGEKTYNPLADFLASTQK 110

QY 121 LYDAKVERDFTNHLNEDTRNINKWVENETHGKIKNVIGEGISSAVMVLVNAVYFKG 180
Db 111 MYGAELASVDFQCAPEDARKEINWVGQTEKIPELLVKGWVDNMTKLVLVNAIYFGN 170

QY 181 WCASTFKSTINCHFKSPKCSKAVAMHQBKRNLSVIEDPSMKILELRNG-GINMYV 239
Db 171 WQKFKRETRDAPRLNKKDTKTVMQYQKKFPYNYIEDLKCRVLELPYQKELSMII 230

QY 240 LLPEN-----DLSIEINKLTFQNLMEWTNPRMTSKYVEVFPQPKIEKNYEMKQVLA 293
Db 231 LLPDIEDESTGLEKLEKQJLTLEKREWTKPNLYLAENVVHLPRFKLEESYDITSELA 290

QY 294 LGUKDIPDSKADLSGIASGRRLYISRMHKSXYIEVTEGTATATATSGNIVEKQLPQST 353
Db 291 LGVQDLFNRGKADLSGSGARDLFVSKTIHKSFVDLNEEGTEAAAAATAGTITLAWLMP 350

QY 354 LFRADHPPLFVIRKDDI--LFPQK 376
Db 351 NFNADHPPIFFIRNPSANILFLGR 375

RESULT 9
US-08-477-108A-8
; Sequence 8, Application US/08477108A
; Patent No. 5801001
;
; GENERAL INFORMATION:
; APPLICANT: Sager, Ruth
; APPLICANT: Zou, Zhigang
; APPLICANT: Anisowicz, Anthony
; TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH
; TITILE OF INVENTION: TUMOR SUPPRESSING ACTIVITY
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,108A
; FILING DATE: June 7, 1995
; CLASSIFICATION: 536

```


TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH TUMOR SUPPRESSING ACTIVITY

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson

STREET: 225 Franklin Street

CITY: Boston

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM PS/2 Model 502 or 55SX

OPERATING SYSTEM: MS-DOS (Version 5.0)

SOFTWARE: WordPerfect (Version 5.1)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/08322

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/938,823

FILING DATE: 09/01/92

APPLICATION NUMBER: 07/844,296

FILING DATE: 02/28/92

APPLICATION NUMBER: 07/662,216

FILING DATE: 02/28/91

ATTORNEY/AGENT INFORMATION:

NAME: Fraser, Janis K.

REGISTRATION NUMBER: 34,819

REFERENCE/DOCKET NUMBER: 00530/072001

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 542-5070

TELEFAX: (617) 542-8906

TELEX: 200154

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 375

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

PCT-US93-08322-8

Query Match 37.3%; Score 730.5; DB 5; Length 375;
Best Local Similarity 37.9%; Pred. No. 2.2e-61;
Matches 146; Conservative 85; Mismatches 135; Indels 19; Gaps 4;

QY 1 MASLAAAEFCFLREEMDDNQGNVFPSSLSLPAALVRLGAQDLSQIDKLHV 60

Db 1 MEQLSTATHPAVDLFRALNESDPTGNIFSPISLSALAMIFLGTGRTAAQVSKALYF 60

QY 61 NTASGYGNSNSQSGLOSKRVFSDINAGHKDYDLSIVNGLFAKVVGFHKDYIECAEK 120

Db 61 DTV-----EDHSRFQSLNADINKPGASYILKANRLYGEKTYNLFADFLASTQK 110

QY 121 LYDAKVERVDTHNLEDTRNINKVNEETHGKIKNVIGEGISSAWVLVNAVYFKG 180

Db 111 MYGAELASVDFQAPEDARKEINWVKQTEGKIPELLVKGWDMTKVLVNAVYFKGN 170

QY 181 WSAFTKSETINCHFKSPCKSKAVAMHGERFNLVSVIEDPSMKILRLYNG-GINMYV 239

Db 171 WQKFMKATNDAPFLNKKOTKTVKMYQKKFAYFYIEDLKRCVLELPYQGEELSMII 230

QY 240 LLPEN-----DLSELENKLTQNLMEWTNPRMTSKVVEFFPQFKIKENYEMQYLRA 293

Db 231 LLPDDIEDSTGLKIEQLTLEKLRWTKPENLYIAEVNHLFRFKLESVDLSEHAR 290

QY 294 LGLKIDFBSKADLSAGGLYIIRNMHKSIVETEGTEATAATGNSNIVEKLPPOST 353

Db 291 LGVQLFNRGKADLSGARDLFVSKIIHKSVDLNEEGTEAAAAATGTMILANLPEE 350

QY 354 LFRADHPFLVIRKDDI--ILFSKG 376

Db 351 NFNAHPPIFFIRHNPANSNIIFLGR 375

RESULT 12

US-08-121-714-4

; Sequence 4, Application US/08121714

; Patent No. 5470970

; GENERAL INFORMATION:

; APPLICANT: Sager, Ruth

; TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH

; TUMOR SUPPRESSING ACTIVITY

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: U.S.A.

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

; COMPUTER: IBM PS/2 Model 502 or 55SX

; OPERATING SYSTEM: MS-DOS (Version 5.0)

; SOFTWARE: WordPerfect (Version 5.1)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/121,714

; FILING DATE:

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/938,823

; FILING DATE: 09/01/92

; APPLICATION NUMBER: 07/844,296

; FILING DATE: 02/28/92

; APPLICATION NUMBER: 07/662,216

; FILING DATE: 02/28/91

; ATTORNEY/AGENT INFORMATION:

; NAME: Fraser, Janis K.

; REGISTRATION NUMBER: 34,819

; REFERENCE/DOCKET NUMBER: 00530/072001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 542-5070

; TELEFAX: (617) 542-8906

; TELEX: 200154

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 379

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; US-08-121-714-4

Query Match 36.9%; Score 722.5; DB 1; Length 379;

Best Local Similarity 38.6%; Pred. No. 1.3e-60;

Matches 150; Conservative 74; Mismatches 146; Indels 19; Gaps 4;

QY 1 MASLAAAEFCFLREEMDDNQGNVFPSSLSLPAALVRLGAQDLSQIDKLHV 60

Db 1 MEQLSANTRALDIFALSENPNPAGNIFSPISLSAMAVFLGTGRTAAQLSTHFP 60

QY 61 NTASGYGNSNSQSGLOSKRVFSDINAGHKDYDLSIVNGLFAKVVGFHKDYIECAEK 120

Db 61 NTV-----EVSHSRFQSLNADINKPGASYILKANRLYGEKTYNLFPEFLVSTQK 110

QY 121 LYDAKVERVDTHNLEDTRNINKVNEETHGKIKNVIGEGISSAWVLVNAVYFKG 180

Db 111 TYGAELASVDFQAPEDARKEINWVKQTEGKIPELLASGWDNMTKVLVNAVYFKGN 170

QY 181 WSAFTKSETINCHFKSPCKSKAVAMHGERFNLVSVIEDPSMKILRLYNG-GINMYV 239

Db 171 WQKFMKATNDAPFLNKKOTKTVKMYQKKFAYFYIEDLKRCVLELPYQGEELSMVI 230

QY 240 LLPEN-----DLSELENKLTQNLMEWTNPRMTSKVVEFFPQFKIKENYEMQYLRA 293

Db 231 LLPDDIEDSTGLKIEQLTLEKLRWTKPENLYIAEVNHLFRFKLESVDLSEHAR 290

QY 294 LGLKDFDESKADLSGASGRILYISRMWHSYIEVTEGTEATAATGCSNIVEKQLPOST 353
 Db 291 LGVQDLFNSKADLSGSGARDIFISKIVHKSFEVNEEGTEAAAAATAGIAFCMLPPE 350
 QY 354 LFRADHPFLFVIRKDD--IILFSGKVSCP 380
 Db 351 NPTADHPFLFIRNSSGSLILFLGRFSSP 379

RESULT 13

US-08-477-108A-4
 ; Sequence 4, Application US/08477108A
 ; Patent No. 5801001
 ; GENERAL INFORMATION:
 ; APPLICANT: Sager, Ruth
 ; APPLICANT: Zou, Zhigang
 ; APPLICANT: Anisowicz, Anthony
 ; TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH
 ; TITLE OF INVENTION: TUMOR SUPPRESSING ACTIVITY
 ; NUMBER OF SEQUENCES: 11
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson P.C.
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: U.S.A.
 ; ZIP: 02110-2804

COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 COMPUTER: IBM PS/2 Model 502 or 55SX
 OPERATING SYSTEM: MS-DOS (Version 5.0)
 SOFTWARE: WordPerfect (Version 5.1)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/477,108A
 FILING DATE: June 7, 1995
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/121,714
 FILING DATE: 09/01/93
 APPLICATION NUMBER: 07/938,823
 FILING DATE: 09/01/92
 APPLICATION NUMBER: 07/844,296
 FILING DATE: 02/28/92
 APPLICATION NUMBER: 07/662,216
 FILING DATE: 02/28/91
 ATTORNEY/AGENT INFORMATION:
 NAME: Fraser, Janis K.
 REGISTRATION NUMBER: 34,819
 REFERENCE/DOCKET NUMBER: 06570/002002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 542-5070
 TELEFAX: (617) 542-8906
 TELEX: 200154
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 379
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear

US-08-477-108A-4

Query Match 36.9%; Score 722.5; DB 1; Length 379;
 Best Local Similarity 38.6%; Pred. No. 1.3e-60;
 Matches 150; Conservative 74; Mismatches 146; Indels 19; Gaps 4;

QY 1 MASLAANAEPFNLFFEMDNGQGNVFFSSLSLFAALVRLGAGDQSLSDIKLLHV 60
 Db 1 MEQLSSANTRALDLFLALSENPNAGNIFISFSSISSAMAVFLGTRGNTAQLSKTHF 60
 QY 61 NTASGYNSSNSQSGLOSKRVFSDINASHKDYLSVNGLFPAKVYGHKDYIECKEK 120
 Db 61 NTV-----BEVHSRFQSNADINKGASYILKANRLYGEKTYNLPFLVSTQK 110

QY 121 LYDAKVERVDFTNHLERTRNINKWVENETHGKIKNIVGEGSISSAVNLVNAVIFKCK 180
 Db 111 TYGADLASVDYFQHASDARKTNQWVKGQTEGKIPELLASGMDVDMTKLVVNAVIFKEN 170
 QY 181 WQSFTKSTETINCHFKSPKSGKAVAMMHQBRKENLSVIEDSPSKILRLYNG-GINMYV 239
 Db 171 WKQKFKKEATNAPFLNKKDKRTYVMYQKKFAYGYIEDLKCRVLELPYQGEELSMTI 230
 QY 240 LLPEN-----DLSEIENKLTQNLMTWNTNPRMTSKYVEVFPQFKIEKNYMKQYLA 293
 Db 231 LLPDDIEDBSTGLKKIEBQJTLKLEHWTKEPDLFIEVNVSLPRFKLEBSYTLNSDLAR 290
 QY 294 LGLKDFDESKADLSGASGRILYISRMWHSYIEVTEGTEATAATGCSNIVEKQLPOST 353
 Db 291 LGVQDLFNSKADLSGSGARDIFISKIVHKSFEVNEEGTEAAAAATAGIAFCMLPPE 350
 QY 354 LFRADHPFLFVIRKDD--IILFSGKVSCP 380
 Db 351 NPTADHPFLFIRNSSGSLILFLGRFSSP 379

RESULT 14

US-08-477-112-4
 ; Sequence 4, Application US/08477112
 ; Patent No. 5905023
 ; GENERAL INFORMATION:
 ; APPLICANT: Sager, Ruth
 ; TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH
 ; TITLE OF INVENTION: TUMOR SUPPRESSING ACTIVITY
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson P.C.
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: U.S.A.
 ; ZIP: 02110-2804

COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 COMPUTER: IBM PS/2 Model 502 or 55SX
 OPERATING SYSTEM: MS-DOS (Version 5.0)
 SOFTWARE: WordPerfect (Version 5.1)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/477,112
 FILING DATE:
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/121,714
 FILING DATE: 09/01/93
 APPLICATION NUMBER: 07/938,823
 FILING DATE: 09/01/92
 APPLICATION NUMBER: 07/844,296
 FILING DATE: 02/28/92
 APPLICATION NUMBER: 07/662,216
 FILING DATE: 02/28/91
 ATTORNEY/AGENT INFORMATION:
 NAME: Fraser, Janis K.
 REGISTRATION NUMBER: 34,819
 REFERENCE/DOCKET NUMBER: 06570/002003
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 542-5070
 TELEFAX: (617) 542-8906
 TELEX: 200154
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 379
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear

US-08-477-112-4

Query Match 36.9%; Score 722.5; DB 2; Length 379;
 Best Local Similarity 38.6%; Pred. No. 1.3e-60;

Matches 150; Conservative 74; Mismatches 146; Indels 19; Gaps 4;

```

QY 1 MASLAANAECFCNLFREDDNOCNGNVFSSLSLFAALALVRLGAODDSLSQIDKILHV 60
Db 1 MEQJSSANTFAIDLFLALSENPNAGNIFISFSSAWAVFLGTRGNTAAQLSKTHF 60
QY 61 NTASGYGNSNSQSLQKRVFSDINASHKDYDLISVINGLFAEKVYGFHDKYIECAEK 120
Db 61 NTV-----EEVHSRFQSLNADINKRGASYILKLANLYGKTYNLFPEFLVSTQK 110
QY 121 LYDAKVERVDFTNHLNEDTRNINKNWVENETHGKINNVIGEGGSSSAVWVAVYFKGK 180
Db 111 TYGADLASVDFOHASEDARKTINQWKGTEGKIPELLASGMVDNMTKILVNAVYFKGN 170
QY 181 WQSAFTKSETINCHPKSPKCGKAVAMHQBKFNLSVIEDPSMKILELYNG-GINMYV 239
Db 171 WKDFKMEATTNAPFLNKKDKRTVMQYCKKFAYGIEDLRCRVLELPYQGBELSMVI 230
QY 240 LLPEN-----DLSEIENKLTFOQLMEWTNPRMTSKYVEVFPFQPKIEKQYEMKQYLRA 293
Db 231 LLPDDIDESGTGLKIEEQLTLEKLEHWTKPENLDFIEVNVSLPRFLEESTYTLNSDLAR 290
QY 294 LGLKDIPEKADSLASGGRLYISMMHKSIVIEVEEGTEATEATGNSNIVERKQLPOST 353
Db 291 LGVQDLFNSKADLSGMSGARDIFISKIVHKSFEVNEEGTEAAAAATAGIATFCWLMPEE 350
QY 354 LFRADHPFLFVIRKDD--IILFSGKYSCP 380
Db 351 NFTAHPFLFFIRHNSGSGSILFLGRFSSP 379

```

RESULT 15

```

PCT-US93-08322-4
; Sequence 4, Application PC/TUS9308322
; GENERAL INFORMATION:
; APPLICANT: Sager, Ruth
; TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH TUMOR SUPPRESSING ACTIVITY
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50z or 55sx
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08322
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/938,823
; FILING DATE: 09/01/92
; APPLICATION NUMBER: 07/844,296
; FILING DATE: 02/26/92
; APPLICATION NUMBER: 07/662,216
; FILING DATE: 02/28/91
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00530/072001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379
; TYPE: amino acid

```

STRANDEDNESS:
TOPOLOGY: linear
PCT-US93-08322-4

Query Match 36.9%; Score 722.5; DB 5; Length 379;
Best Local Similarity 38.6%; Pred. No. 1.3e-60;
Matches 150; Conservative 74; Mismatches 146; Indels 19; Gaps 4;

```

QY 1 MASLAANAECFCNLFREDDNOCNGNVFSSLSLFAALALVRLGAODDSLSQIDKILHV 60
Db 1 MEQJSSANTFAIDLFLALSENPNAGNIFISFSSAWAVFLGTRGNTAAQLSKTHF 60
QY 61 NTASGYGNSNSQSLQKRVFSDINASHKDYDLISVINGLFAEKVYGFHDKYIECAEK 120
Db 61 NTV-----EEVHSRFQSLNADINKRGASYILKLANLYGKTYNLFPEFLVSTQK 110
QY 121 LYDAKVERVDFTNHLNEDTRNINKNWVENETHGKINNVIGEGGSSSAVWVAVYFKGK 180
Db 111 TYGADLASVDFOHASEDARKTINQWKGTEGKIPELLASGMVDNMTKILVNAVYFKGN 170
QY 181 WQSAFTKSETINCHPKSPKCGKAVAMHQBKFNLSVIEDPSMKILELYNG-GINMYV 239
Db 171 WKDFKMEATTNAPFLNKKDKRTVMQYCKKFAYGIEDLRCRVLELPYQGBELSMVI 230
QY 240 LLPEN-----DLSEIENKLTFOQLMEWTNPRMTSKYVEVFPFQPKIEKQYEMKQYLRA 293
Db 231 LLPDDIDESGTGLKIEEQLTLEKLEHWTKPENLDFIEVNVSLPRFLEESTYTLNSDLAR 290
QY 294 LGLKDIPEKADSLASGGRLYISMMHKSIVIEVEEGTEATEATGNSNIVERKQLPOST 353
Db 291 LGVQDLFNSKADLSGMSGARDIFISKIVHKSFEVNEEGTEAAAAATAGIATFCWLMPEE 350
QY 354 LFRADHPFLFVIRKDD--IILFSGKYSCP 380
Db 351 NFTAHPFLFFIRHNSGSGSILFLGRFSSP 379

```

Search completed: December 12, 2003, 16:20:06
Job time : 23 secs


```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-067339
; FILING DATE: 04-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-247
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-140-719-34

```

```

Query Match 100.0%; Score 1959; DB 9; Length 380;
Best Local Similarity 100.0%; Pred. No. 8e-172;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MASLAANAARFCNLFREMDNQGNGVFPSSLSLFAALALVRLGACDSSLSQIDKLLHV 60
DB 1 MASLAANAARFCNLFREMDNQGNGVFPSSLSLFAALALVRLGACDSSLSQIDKLLHV 60
QY 61 NTASGYGNSNSOSGLQSLKRVFSDINASHKDYDLSTVNGLPFAEKVYGFHKDYIECAEK 120
DB 61 NTASGYGNSNSOSGLQSLKRVFSDINASHKDYDLSTVNGLPFAEKVYGFHKDYIECAEK 120
QY 121 LYDAKVRVDFTNHLEDTRRNINKNWENETHGKIKNNVIGEGGSSAVMVLNAVYFKG 180
DB 121 LYDAKVRVDFTNHLEDTRRNINKNWENETHGKIKNNVIGEGGSSAVMVLNAVYFKG 180
QY 181 WSAFTKSETINCHFKSPKCSGKAVAMHQRKPNLSVIEDPSMKILELYNGGINMYVL 240
DB 181 WSAFTKSETINCHFKSPKCSGKAVAMHQRKPNLSVIEDPSMKILELYNGGINMYVL 240
QY 241 LPENDLSIEINKLTFFQNLMEWTNPRMTSKYVEVFPFQFKIEKNYEMKQYLRALGLKDF 300
DB 241 LPENDLSIEINKLTFFQNLMEWTNPRMTSKYVEVFPFQFKIEKNYEMKQYLRALGLKDF 300
QY 301 DESKADLSGIASGRLYISRMMHKSYLEVTEEGTATATGNSNIVEKQLPQSTLFRADHP 360
DB 301 DESKADLSGIASGRLYISRMMHKSYLEVTEEGTATATGNSNIVEKQLPQSTLFRADHP 360
QY 361 FLFVIRKDDIILFSGKVSCP 380
DB 361 FLFVIRKDDIILFSGKVSCP 380

```

RESULT 2

US-10-091-442-34

; Sequence 34, Application US/10091442

; Publication No. US20020164711A1

; GENERAL INFORMATION:

; APPLICANT: TSUJIMOTO, Masafumi

; IWASAKI, Fuyuki

; TSURUOKA, No. US20020164711A1

; NAKAZATO, Hiroshi

; MURA, Kenju

; ISHIDA, No. US20020164711A1

; KURIHARA, Tatsuya

; YAMAUCHI, Kozo

; YAMAGUCHI, No. US20020164711A1

; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR

; NUMBER OF SEQUENCES: 34

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Burns, Doane, Swecker & Mathis

; STREET: P.O. Box 1404

```

; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/091,442
; FILING DATE: 07-Mar-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/140,719
; FILING DATE: 08-AUG-1998
; APPLICATION NUMBER: US 08/474,661
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; APPLICATION NUMBER: JP 6-067339
; FILING DATE: 04-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-247
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-10-091-442-34

```

```

Query Match 100.0%; Score 1959; DB 14; Length 380;
Best Local Similarity 100.0%; Pred. No. 8e-172;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASLAANAARFCNLFREMDNQGNGVFPSSLSLFAALALVRLGACDSSLSQIDKLLHV 60
DB 1 MASLAANAARFCNLFREMDNQGNGVFPSSLSLFAALALVRLGACDSSLSQIDKLLHV 60
QY 61 NTASGYGNSNSOSGLQSLKRVFSDINASHKDYDLSTVNGLPFAEKVYGFHKDYIECAEK 120
DB 61 NTASGYGNSNSOSGLQSLKRVFSDINASHKDYDLSTVNGLPFAEKVYGFHKDYIECAEK 120
QY 121 LYDAKVRVDFTNHLEDTRRNINKNWENETHGKIKNNVIGEGGSSAVMVLNAVYFKG 180
DB 121 LYDAKVRVDFTNHLEDTRRNINKNWENETHGKIKNNVIGEGGSSAVMVLNAVYFKG 180
QY 181 WSAFTKSETINCHFKSPKCSGKAVAMHQRKPNLSVIEDPSMKILELYNGGINMYVL 240
DB 181 WSAFTKSETINCHFKSPKCSGKAVAMHQRKPNLSVIEDPSMKILELYNGGINMYVL 240
QY 241 LPENDLSIEINKLTFFQNLMEWTNPRMTSKYVEVFPFQFKIEKNYEMKQYLRALGLKDF 300
DB 241 LPENDLSIEINKLTFFQNLMEWTNPRMTSKYVEVFPFQFKIEKNYEMKQYLRALGLKDF 300
QY 301 DESKADLSGIASGRLYISRMMHKSYLEVTEEGTATATGNSNIVEKQLPQSTLFRADHP 360
DB 301 DESKADLSGIASGRLYISRMMHKSYLEVTEEGTATATGNSNIVEKQLPQSTLFRADHP 360
QY 361 FLFVIRKDDIILFSGKVSCP 380
DB 361 FLFVIRKDDIILFSGKVSCP 380

```

RESULT 3

US-10-198-070-62

; Sequence 62, Application US/10198070

; Publication No. US20030109437A1

; GENERAL INFORMATION:

; APPLICANT: AVERBACK, PAUL

; APPLICANT: GEMWELL, JACK

; TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER

; TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF

; TITLE OF INVENTION: CELLS

; FILE REFERENCE: 59003.000008

; CURRENT APPLICATION NUMBER: US/10/198,070

; CURRENT FILING DATE: 2002-07-19

; PRIOR APPLICATION NUMBER: 60/306,161

; PRIOR FILING DATE: 2001-07-19

; PRIOR APPLICATION NUMBER: 60/306,150

; PRIOR FILING DATE: 2001-07-19

; PRIOR APPLICATION NUMBER: 60/331,477

; PRIOR FILING DATE: 2001-11-16

; NUMBER OF SEQ ID NOS: 125

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 62

; LENGTH: 392

; TYPE: PRP

; ORGANISM: Homo sapiens

US-10-198-070-62

Query Match

Best Local Similarity 42.0%; Score 822; DB 15; Length 392;

Mismatches 169; Conservative 80; Mismatches 131; Indels 12; Gaps 6;

QY 1 MASLAAANAFPCNLFREMDNQNGNVFFSSLSLFAALALVRLGAQDLSQIDKLLHV 60

DB 1 MGSLSANVEFCLDVFKEINNNIGDNIFSSLSLLYALSWLLGARGETAQLEKVLHF 60

QY 61 -----NTASGYGNSN--SQSG--LOSQKRVFSDINASHKDYDLSIVNGLFAKVGREKD 113

DB 61 SHTVDSLKPGFKDPSKCSQAGRIHSEFGVFSQINQPNCTLSIANRLYGTKTWAFHQ 120

QY 114 YIECAEKLYDAKVERVDFTNHLETRRNINKWVENETHGKIKNVIGEGISSAVMVLN 173

DB 121 YLSCSEKWKQARLQTVDFEQSTETRTKINAWENKINGKVANLFGKSTIDPSVWVLN 180

QY 174 AVYFKGKQSAFTKSTINCHFKSPKCSGKAVAMQERKFNLSVIEDPSMKILRLY-N 232

DB 181 TIYFKGRQKRFQRTVKSPFQSGKQNTVEMMYQIGTKLAFVKEPQWVLELPYN 240

QY 233 GGINMYVLLPEN--DLSEIENKLTQNLMEWTNPRMTSKYVEVFPQKIEKNYEMKQY 290

DB 241 NKLSMILLPVGIANLKQIEKQNSGTFHEWTSNNMREVEVHLPRFKLBKIELNSL 300

QY 291 LRALGLKDIFDESKADLSGAGGRILYISRMHKSIVTEBGTETAATGNSNIVEKQLP 350

DB 301 LKPLGVTLDFNQYKADLSGMSFTKGLYLSKAIHKSILYDVSSEGTAAATGDSIAVKSLP 360

QY 351 QSTLFRADHPFLFVIR--KDDIILFSGKVSCP 380

DB 361 MEAQFKANHPFLFIRHTHTNTILFCGLASP 392

RESULT 4

US-10-198-070-67

; Sequence 67, Application US/10198070

; Publication No. US20030109437A1

; GENERAL INFORMATION:

; APPLICANT: AVERBACK, PAUL

; APPLICANT: GEMWELL, JACK

; TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER

; TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF

; TITLE OF INVENTION: CELLS

; FILE REFERENCE: 59003.000008

; CURRENT APPLICATION NUMBER: US/10/198,070

; CURRENT FILING DATE: 2002-07-19

; QUERY MATCH

Best Local Similarity 42.0%; Score 822; DB 15; Length 392;

Mismatches 169; Conservative 80; Mismatches 131; Indels 12; Gaps 6;

QY 1 MASLAAANAFPCNLFREMDNQNGNVFFSSLSLFAALALVRLGAQDLSQIDKLLHV 60

DB 1 MGSLSANVEFCLDVFKEINNNIGDNIFSSLSLLYALSWLLGARGETAQLEKVLHF 60

QY 61 -----NTASGYGNSN--SQSG--LOSQKRVFSDINASHKDYDLSIVNGLFAKVGREKD 113

DB 61 SHTVDSLKPGFKDPSKCSQAGRIHSEFGVFSQINQPNCTLSIANRLYGTKTWAFHQ 120

QY 114 YIECAEKLYDAKVERVDFTNHLETRRNINKWVENETHGKIKNVIGEGISSAVMVLN 173

DB 121 YLSCSEKWKQARLQTVDFEQSTETRTKINAWENKINGKVANLFGKSTIDPSVWVLN 180

QY 174 AVYFKGKQSAFTKSTINCHFKSPKCSGKAVAMQERKFNLSVIEDPSMKILRLY-N 232

DB 181 TIYFKGRQKRFQRTVKSPFQSGKQNTVEMMYQIGTKLAFVKEPQWVLELPYN 240

QY 233 GGINMYVLLPEN--DLSEIENKLTQNLMEWTNPRMTSKYVEVFPQKIEKNYEMKQY 290

DB 241 NKLSMILLPVGIANLKQIEKQNSGTFHEWTSNNMREVEVHLPRFKLBKIELNSL 300

QY 291 LRALGLKDIFDESKADLSGAGGRILYISRMHKSIVTEBGTETAATGNSNIVEKQLP 350

DB 301 LKPLGVTLDFNQYKADLSGMSFTKGLYLSKAIHKSILYDVSSEGTAAATGDSIAVKSLP 360

QY 351 QSTLFRADHPFLFVIR--KDDIILFSGKVSCP 380

DB 361 MEAQFKANHPFLFIRHTHTNTILFCGLASP 392

; PRIOR APPLICATION NUMBER: 60/306,161

; PRIOR FILING DATE: 2001-07-19

; PRIOR APPLICATION NUMBER: 60/306,150

; PRIOR FILING DATE: 2001-07-19

; PRIOR APPLICATION NUMBER: 60/331,477

; PRIOR FILING DATE: 2001-11-16

; NUMBER OF SEQ ID NOS: 125

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 67

; LENGTH: 392

; TYPE: PRP

; ORGANISM: Homo sapiens

US-10-198-070-67

Query Match

Best Local Similarity 41.9%; Score 820; DB 15; Length 392;

Mismatches 169; Conservative 80; Mismatches 131; Indels 12; Gaps 6;

QY 1 MASLAAANAFPCNLFREMDNQNGNVFFSSLSLFAALALVRLGAQDLSQIDKLLHV 60

DB 1 MGSLSANVEFCLDVFKEINNNIGDNIFSSLSLLYALSWLLGARGETAQLEKVLHF 60

QY 61 -----NTASGYGNSN--SQSG--LOSQKRVFSDINASHKDYDLSIVNGLFAKVGREKD 113

DB 61 SHTVDSLKPGFKDPSKCSQAGRIHSEFGVFSQINQPNCTLSIANRLYGTKTWAFHQ 120

QY 114 YIECAEKLYDAKVERVDFTNHLETRRNINKWVENETHGKIKNVIGEGISSAVMVLN 173

DB 121 YLSCSEKWKQARLQTVDFEQSTETRTKINAWENKINGKVANLFGKSTIDPSVWVLN 180

QY 174 AVYFKGKQSAFTKSTINCHFKSPKCSGKAVAMQERKFNLSVIEDPSMKILRLY-N 232

DB 181 TIYFKGRQKRFQRTVKSPFQSGKQNTVEMMYQIGTKLAFVKEPQWVLELPYN 240

QY 233 GGINMYVLLPEN--DLSEIENKLTQNLMEWTNPRMTSKYVEVFPQKIEKNYEMKQY 290

DB 241 NKLSMILLPVGIANLKQIEKQNSGTFHEWTSNNMREVEVHLPRFKLBKIELNSL 300

QY 291 LRALGLKDIFDESKADLSGAGGRILYISRMHKSIVTEBGTETAATGNSNIVEKQLP 350

DB 301 LKPLGVTLDFNQYKADLSGMSFTKGLYLSKAIHKSILYDVSSEGTAAATGDSIAVKSLP 360

QY 351 QSTLFRADHPFLFVIR--KDDIILFSGKVSCP 380

DB 361 MEAQFKANHPFLFIRHTHTNTILFCGLASP 392

RESULT 5

US-10-198-070-70

; Sequence 70, Application US/10198070

; Publication No. US20030109437A1

; GENERAL INFORMATION:

; APPLICANT: AVERBACK, PAUL

; APPLICANT: GEMWELL, JACK

; TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER

; TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF

; TITLE OF INVENTION: CELLS

; FILE REFERENCE: 59003.000008

; CURRENT APPLICATION NUMBER: US/10/198,070

; CURRENT FILING DATE: 2002-07-19

; PRIOR APPLICATION NUMBER: 60/306,161

; PRIOR FILING DATE: 2001-07-19

; PRIOR APPLICATION NUMBER: 60/306,150

; PRIOR FILING DATE: 2001-07-19

; PRIOR APPLICATION NUMBER: 60/331,477

; PRIOR FILING DATE: 2001-11-16

; NUMBER OF SEQ ID NOS: 125

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 70

; LENGTH: 392

; TYPE: PRP

; ORGANISM: Homo sapiens

US-10-198-070-70

Query Match 41.8%; Score 819; DB 15; Length 392;
Best Local Similarity 43.1%; Pred. No. 7.1e-67;
Matches 169; Conservative 80; Mismatches 131; Indels 12; Gaps 6;

```
Qy 1 MASLAANAABFCENLFREMDNNGNGVFFSSLSLFAALALVELGAODDSLSQIDKLLHV 60
Db 1 MGSLSANVEFCLDVFKELSNNGNIFFSSLSLLVALSNVILGARGETAQLEKVIHF 60

Qy 61 -----NTASGYGNSN--SQSG--LQSOLKEVFDINASHKDYDLSIVNGLFAEKYVGFHD 113
Db 61 SHTVDSLKPGFKDSCQAGRIHSFGVPSQINQPSNCTLSIANRLYGTKTMAFHQ 120

Qy 114 YIECAEKLYDAKVERVDTNHLNINQKVENETHGKIKVIGEGISSAVMYLVN 173
Db 121 YLSCEKWKYQARLQTVDFEQSTETRMNINAVENKINGVANLFGKSTIDPSSVMVLN 180

Qy 174 AVYFKGKWSAQTKSETINCHFKSPKCSGKAVAMHQRKPNLSVIEDPSMKILELY-N 232
Db 181 IIVFKGQKQVRETQVRSFQI-SGKNVTVMYQIGTKLAFVKFQVQVLEFPYN 240

Qy 233 GGINMYVLLPEN--DLSEIENKLTFFQNLMEWTNPRMTSKYVEVPPFPQFKIEKNYMKQV 290
Db 241 NKLSMILLVGVIANLAKIQKLSNGTFHEWTSNNMWEREVEVHLPRFKLEIKYELNSL 300

Qy 291 LRALGLKDIPDSKADLSIASGGRLYISRMHKSIVETEGTEATATGNSIVKQLP 350
Db 301 LKPLGVDTLNFQVKADLSGMSPTKGLYLSKAIHKSILVDSEEGTEAATAATGDSIANVSLP 360

Qy 351 QSTLFRADHPELFVIR--KDDIILFGKVSQCP 380
Db 361 MEAQFKANHPFLFFIRHTHTNTILFCGKLASP 392
```

RESULT 6

```
US-09-895-814-946
; Sequence 946, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 946
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-895-814-946
```

Query Match 40.7%; Score 796.5; DB 10; Length 379;
Best Local Similarity 42.6%; Pred. No. 7.9e-65;
Matches 168; Conservative 77; Mismatches 120; Indels 29; Gaps 8;

```
Qy 1 MASLAANAABFCENLFREMDNNGNGVFFSSLSLFAALALVELGAODDSLSQIDKLLHV 60
Db 1 MGSLSANVEFCLDVFKELSNNGNIFFSSLSLLVALSNVILGARGETAQLEKVIHF 60

Qy 61 -----NTASGYGNSN--SQSG--LQSOLKEVFDINASHKDYDLSIVNGLFAEKYVGFHD 113
Db 61 SHTVDSLKPGFKDSCQAGRIHSFGVPSQINQPSNCTLSIANRLYGTKTMAFHQ 120

Qy 114 YIECAEKLYDAKVERVDTNHLNINQKVENETHGKIKVIGEGISSAVMYLVN 173
Db 121 YLSCEKWKYQARLQTVDFEQSTETRMNINAVENKINGVANLFGKSTIDPSSVMVLN 180

Qy 174 AVYFKGKWSAQTKSETINCHFKSPKCSGK--AVAMHQRKPNLSVIEDPSMKILELY 231
Db 181 IIVFKGQKQVRETQVRSFQI-SGKNVTVMYQIGTKLAFVKFQVQVLEFPY 225

Qy 232 GGINMYVLLPEN--DLSEIENKLTFFQNLMEWTNPRMTSKYVEVPPFPQFKIEKNYEMK 288
Db 226 YNNKLSMILLVGVIANLAKIQKLSNGTFHEWTSNNMWEREVEVHLPRFKLEIKYELN 285

Qy 289 QYLRALGLKDIPDSKADLSIASGGRLYISRMHKSIVETEGTEATATGNSIVKQ 348
Db 286 SLKPLGVDTLNFQVKADLSGMSPTKGLYLSKAIHKSILVDSEEGTEAATAATGDSIAVKS 345

Qy 349 LPQSTLFRADHPELFVIR--KDDIILFGKVSQCP 380
Db 346 LPMRAQFKANHPFLFFIRHTHTNTILFCGKLASP 379
```

RESULT 7

```
US-10-144-678A-946
; Sequence 946, Application US/10144678A
; Publication No. US20030157089A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Hepler, William T.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa M.
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Deng, Ta
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C28
; CURRENT APPLICATION NUMBER: US/10/144,678A
; CURRENT FILING DATE: 2002-08-12
; NUMBER OF SEQ ID NOS: 1033
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 946
; LENGTH: 379
; TYPE: PRT
```



```
/ ORGANISM: Homo sapiens
US-10-144-678A-946

Query Match      40.7%; Score 796.5; DB 12; Length 379;
Best Local Similarity 42.6%; Pred. No. 7.9e-65;
Matches 168; Conservative 77; Mismatches 120; Indels 29; Gaps 8;

QY 1 MASLAANAFCNLFREDDNQGNGVFFSSLSFAALALVRLGAQDLSQIDKLHV 60
DQ 1 MGSUSTANVFCLDVFKELSNNGDNIFFSSLSLYALSVMVLGARGTAEQLEKVLHF 60
QY 61 -----NTASGVGNSSN-SQSG-LOSQKRVFSDINASHKDYDLSIVNGLFAKVGPHKD 113
DQ 61 SHTVDSLKPGFDSKPSQAGRIHSEGFVFSQINQPSNCTLSIANRLYGTWAFHQ 120
QY 114 YIECAKLYDAKVERVDFNHLNEDTRNINKWNETGKTKNVI GEGGISSAVWVLN 173
DQ 121 YLSCSEKWOARLQTVDFEQSTETTRKMINAWENKTKGVANLFGKSTIDPSSWVLN 180
QY 174 AVYFKGWSAFTKSTINCHFKSPKCSGK--AVAMHQRKFNLSVIEDPSMKILELY 231
DQ 181 TIYFKGQRQNF-----QGNVTVMYQIGTFKLAFFVKPQWVLELPY 225
QY 232 -NGGINMYVLLPEN--DLSEIKNLTFTONLWMTNPRMTSKYVEFFPOPKIEKNYEMK 288
DQ 226 VNNKLSMILLPVGIANLQIEKQNSGTHETWSSNMWEREVEVHLPRFKLEIKYELN 285
QY 289 QYLRLGKIDFDESADLSGIASGRLYISRMHKSIVYETEGTEATAATGSNIVEKQ 348
DQ 286 SLLKPLGVTDLFNQVADLSGMSPTKGLYLSKAIHKSILYDSEEGTEAAATGDSIAVKS 345
QY 349 LPQSLFRAADHPFLVIR--KDDILFSGKVSCP 380
DQ 346 LPMRAQKXANHPFLFFIRHTHTNLTILFCGLASP 379

RESULT 8
US-10-294-025-946
; Sequence 946, Application US/10294025
; Publication No. US20030185830A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Stolk, Michael D.
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C29
; CURRENT APPLICATION NUMBER: US/10/294.025
; CURRENT FILING DATE: 2002-11-12
; NUMBER OF SEQ ID NOS: 1038
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 946
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-294-025-946

Query Match      40.7%; Score 796.5; DB 12; Length 379;
Best Local Similarity 42.6%; Pred. No. 7.9e-65;
Matches 168; Conservative 77; Mismatches 120; Indels 29; Gaps 8;

QY 1 MASLAANAFCNLFREDDNQGNGVFFSSLSFAALALVRLGAQDLSQIDKLHV 60
DQ 1 MGSUSTANVFCLDVFKELSNNGDNIFFSSLSLYALSVMVLGARGTAEQLEKVLHF 60
QY 61 -----NTASGVGNSSN-SQSG-LOSQKRVFSDINASHKDYDLSIVNGLFAKVGPHKD 113
DQ 61 SHTVDSLKPGFDSKPSQAGRIHSEGFVFSQINQPSNCTLSIANRLYGTWAFHQ 120
QY 114 YIECAKLYDAKVERVDFNHLNEDTRNINKWNETGKTKNVI GEGGISSAVWVLN 173
DQ 121 YLSCSEKWOARLQTVDFEQSTETTRKMINAWENKTKGVANLFGKSTIDPSSWVLN 180
QY 174 AVYFKGWSAFTKSTINCHFKSPKCSGK--AVAMHQRKFNLSVIEDPSMKILELY 231
DQ 181 TIYFKGQRQNF-----QGNVTVMYQIGTFKLAFFVKPQWVLELPY 225
QY 232 -NGGINMYVLLPEN--DLSEIKNLTFTONLWMTNPRMTSKYVEFFPOPKIEKNYEMK 288
DQ 226 VNNKLSMILLPVGIANLQIEKQNSGTHETWSSNMWEREVEVHLPRFKLEIKYELN 285
QY 289 QYLRLGKIDFDESADLSGIASGRLYISRMHKSIVYETEGTEATAATGSNIVEKQ 348
DQ 286 SLLKPLGVTDLFNQVADLSGMSPTKGLYLSKAIHKSILYDSEEGTEAAATGDSIAVKS 345
QY 349 LPQSLFRAADHPFLVIR--KDDILFSGKVSCP 380
DQ 346 LPMRAQKXANHPFLFFIRHTHTNLTILFCGLASP 379

RESULT 9
US-10-012-896-946
; Sequence 946, Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yudi
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Keppler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Rural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carola
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; APPLICANT: Katanabe, Yoshihiro
; APPLICANT: Zeagher, Madeleine Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C27
; CURRENT APPLICATION NUMBER: US/10/012.896
; CURRENT FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 1011
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 946
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-012-896-946

Query Match      40.7%; Score 796.5; DB 14; Length 379;
Best Local Similarity 42.6%; Pred. No. 7.9e-65;
Matches 168; Conservative 77; Mismatches 120; Indels 29; Gaps 8;

QY 1 MASLAANAFCNLFREDDNQGNGVFFSSLSFAALALVRLGAQDLSQIDKLHV 60
DQ 1 MGSUSTANVFCLDVFKELSNNGDNIFFSSLSLYALSVMVLGARGTAEQLEKVLHF 60
QY 61 -----NTASGVGNSSN-SQSG-LOSQKRVFSDINASHKDYDLSIVNGLFAKVGPHKD 113
DQ 61 SHTVDSLKPGFDSKPSQAGRIHSEGFVFSQINQPSNCTLSIANRLYGTWAFHQ 120
QY 114 YIECAKLYDAKVERVDFNHLNEDTRNINKWNETGKTKNVI GEGGISSAVWVLN 173
DQ 121 YLSCSEKWOARLQTVDFEQSTETTRKMINAWENKTKGVANLFGKSTIDPSSWVLN 180
```

Db 121 YLSCSEKWKQARLOTVDPEQSTBETRKMINAWENKTNCKVANLFGKSTIDPSSVMVLVN 180
Qy 174 AVYFKGWOSAFKSEIINCHFKSPKCSK--AVAMMEQRKFNLSVIEDPSMKILELRY 231
Db 181 TIFPKQQRNK-----QCKNTVEMMYQIGTFKLAFKVEFQMOVLELRY 225
Qy 232 -NGGINMYLLPEN--DLSEIENKLTFFQNAEWNTNPRMTSKYVFPFQFKIKENYEMK 288
Db 256 VNNKLSMIILLPVGIANLKQIEKQLNSGTFHEWTSNMMEREVEVHLPRFKLEIKYELN 285
Qy 289 QYLRLALGLKIDFDSKADLSGIASGGRLYISRMHKSIVYTBEGTEATATGSNIVEKQ 348
Db 286 SLUKPLGVLDLNFQWAKDLGSGMFTKGLYLKAIHKSYLDVSEEGTEAARATGSAVKS 345
Qy 348 LPQSTLFRADHPFLFVIR--KDDIILFSGKVSCLP 380
Db 346 LPKRAQKANHFFLFFIRHHTWILLFCGLKASLP 379

RESULT 10
US-09-823-187-2
; Sequence 2, Application US/09823187
; Publication No. US2003009652A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Catherine
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Liu, Xiaohong
; APPLICANT: Majumder, Kumud
; APPLICANT: Padigaru, Muralidhar
; APPLICANT: Patturajan, Meera
; APPLICANT: Shukrets, Richard A
; APPLICANT: Spaderna, Steven K
; APPLICANT: Spytek, Kimberly
; APPLICANT: Taudier, Raymond J
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-745
; CURRENT APPLICATION NUMBER: US/09/823,187
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/193,339
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/193,205
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/195,343
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/195,088
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,005
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,792
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: 60/196,556
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: 60/197,081
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/197,525
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/197,087
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-823-187-2

Query Match 39.4%; Score 771.5; DB 11; Length 395;
Best Local Similarity 40.3%; Pred. No. 1.7e-62;
Matches 159; Conservative 89; Mismatches 132; Indels 15; Gaps 5;
Qy 1 MASLAANAFCNLFREDDNQNGNVFSSLSFAALALVRLGAQDDSLQIDKLHV 60
Db 1 MDSLVTANTKCFDLFQEIQKDRHKNIFFSPLSLSAALGMVRLGARSASAHQIDVLHF 60

Qy 61 NTASGY-----GNSSNSQSGLOS-OLKRVFSDINASHKDYDLSIVNGLFAEKVYGFHKDY 114
Db 61 NEFQNSKEPAGSLNNESGLVSCYFQQLLSKLDRIKTDYTLSTIANRLYGEFFIQEY 120
Qy 115 IECSEKLYDAKVERVDPNHLEDTNRINKAVENETHGKIKNVIGGGISSATWVLVNA 174
Db 121 LDGVIQYFHTTESVDFOKPEKSQRIENFWVECOSQKIKDLFSKDAINAEIVLVLNA 180
Qy 175 VYFKGWQSAFTKSTINCHFKSPKCSKAVAMHQRKFNLSVIDPMSKILELRY-NG 233
Db 181 VYFAKWEIYFHEHTVDPAPCLNQNENKSVQMTQKGLYRIGFIEVKAQILEMRYTKG 240
Qy 234 GINMYVLLPND-----LSIENKLTFFQNAEWNTNPRMTSKYVFPFQFKIKENYEM 287
Db 241 KLSMFVILLPSHSGKDLGLLELERKITYEKMWANSSSENMBESVVLSPRFTLESDYL 300
Qy 288 QYLRLALGLKIDFDSKADLSGIASGGRLYISRMHKSIVYTBEGTEATATGSNIVEK 347
Db 301 NSIILQMGITDIFDETRADLTGISPSNLYLSKIHTFVEVDENGTOAATGAVVSR 360
Qy 348 QLPQSTLFRADHPFLFVIR--KDDIILFSGKVSCLP 380
Db 361 SLRSWVEFNHFFLFFIRHNTQTILFYGRVCSLP 395

RESULT 11
US-10-094-944-2
; Sequence 2, Application US/10094944
; Publication No. US20030150275A1
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Clarke, Howard RG
; APPLICANT: DuBoise, Robert F
; APPLICANT: Wiley, Steven R
; TITLE OF INVENTION: HUMAN SERPIN POLYPEPTIDES
; FILE REFERENCE: 3223-A
; CURRENT APPLICATION NUMBER: US/10/094,944
; CURRENT FILING DATE: 2002-03-08
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 2
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-944-2

Query Match 38.5%; Score 754.5; DB 12; Length 425;
Best Local Similarity 37.4%; Pred. No. 6.9e-61;
Matches 159; Conservative 87; Mismatches 134; Indels 45; Gaps 5;
Qy 1 MASLAANAFCNLFREDDNQNGNVFSSLSFAALALVRLGAQDDSLQIDKLHV 60
Db 1 MDSLVTANTKCFDLFQEIQKDRHKNIFFSPLSLSAALGMVRLGARSASAHQIDVLHF 60
Qy 61 NTAS-----GFGNSSNSQSGLOS-OLKRVF 84
Db 61 NEFQNSKEPAGSLNNESGLVSCYFQQLLSKLDRIKTDYTLSTIANRLYGEFFIQEY 120
Qy 85 SDINASHKDYDLSIVNGLFAEKVYGFHKDYIECAEKLYDAKVERVDFTNHLDTNRINK 144
Db 121 SKLDRIKTDYTLSTIANRLYGEFFIQEYVLDGVTFYHTIESVDFOKPEKSQRIEN 180
Qy 145 WYENETHGKIKNVIGGGISSAVVNLVNAVYFKGWQSAFTKSTINCHFKSPKCSKA 204
Db 181 WVECSQSKIKELFSKDAINAEIVLVLNAVYFKAKWTYFDHENTVDAPFCINANENK 240
Qy 205 VAMMEQRKFNLSVIDPMSKILELRY-NGGINMYVLLPND-----LSEIENKLTQ 257
Db 241 VRQMTQKLYRIGFIEVKAQILEMRYTKGLSMFVLLPSHSGKDLGLLELERKITYEK 300
Qy 258 LMEWNTNPRMTSKYVFPFQFKIKENYEMKQYLRLALGLKIDFDSKADLSGIASGGRLY 317

```

Db          301  MVAWSSNMSESVLSPRTLESDSDVLSNLTQDMGITHIDFETRADLTGIGSPSNLY 360
QY          318  ISRMHKSXYIEETEGTEATAATGNSNIVEKQLFQSTFRADHPFLFVIR--KDSIIILFSG 375
          ::::||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db          361  LSKIIHKTVEVDENGQTQAAATGAVVSESSLFWSVFNANHPFLFIHKNKTQILFYG 420
QY          376  KVSCP 380
          :|||
Db          421  RVCSP 425

RESULT 12
US-09-823-187-29
; Sequence 29, Application US/09823:187
; Publication No. US20030096952A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Catherine
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Liu, Xiaohong
; APPLICANT: Majumder, Kumud
; APPLICANT: Padiguru, Muralidhar
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard A
; APPLICANT: Spaderna, Steven K
; APPLICANT: Spytek, Kimberly
; APPLICANT: Taupier, Raymond J
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-745
; CURRENT APPLICATION NUMBER: US/09/823:187
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/193,339
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/193,205
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/195,343
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/195,088
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,005
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,792
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: 60/196,556
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: 60/197,081
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/197,525
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/197,087
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-823-187-29

```

Qy	175	VYFKGWQSAFTKSETINCHPKSPKCSGKAVAMHQBERKNLSVIEDPMKILELRYPNG-	233
Db	180	IYFKGWKKENKEDYKEBKFPNKNTYKSIQMRYQTSFZASLEDEQAUKVLPIYKKG	239
Qy	234	GIMNYVLLPE--NDLGSIEKNLTIFONLMWTPRMTSKYVEVPFOPKIKNEMEQYL	291
Db	240	DLSMIVLVLPNEIDLOKLBEKLIKAELMEWTSIONRWTRVDLHLFRFKVESYDLKDTL	299
Qy	292	RALGHAKDITFESKADUSGTAGSGRLYISRMKHKSIVETBEGTEATAAT----	346
Db	300	RTMGWVDIFN-GDAIDLSGTSGRGVLISGLVKHAFVETEERGAATAAUVAFGSS---	355
Qy	347	KQLPOST-LFRADHPFLFVIR--KDDIIILFGSKVSCP	380
Db	356	--PSTNBEFHCHPFLFIQNKTNSILFYGRFSPP	390
RESULT 13			
US-09-823-187-30			

```

Query Match      35.4%; Score 752, DB 11; Length 390;
Best Local Similarity 41.5%; Pred. No. te-60;
Matches 165; Conservative 81; Mismatches 126; Indels 26; Gaps 10;

QY 1 MASLAAANAFCCNFREMDNQNQNGNVFFSLFSLFALALVRLGAQDSDLSIQDKLLHV 60
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MASLSEANTVFDFQFQFRKSEN NIYSPTISITGALGWLLGAKONTAQQIKKLVHF 59
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```
QY 61 -----NTASGYGNSNSQSG-LOSQKRVFSDINASHKDYDLSIVNGLFAEKVYGFHKDY 114
Db 60 DQVTEINTGKATYHVDSEGNVHHQFQKLLIEFNKSTDAVELKIANKLFGEKTYLFQY 119
QY 115 IECAEKLYDAKVERVDFTNHLIEDTRNNKXVETHGKIKNVIGEGISSAVMVLVNA 174
Db 120 LDAIKFYQTSVESVDANAPESRKKINSWESQTFNEKIKNLPEGNIGSNTTLVLVNA 179
QY 175 VYFKGWQSAFTKSTINCHFKPKSGKAVAMHQRKENLSVIEDPSMKILELRNG- 233
Db 180 IYFKGWQKXKFKNEKTEKRFWPNKNTYKSIQMMKQYTSFHFASLEDVQAKVLEIPYKX 239
QY 234 GINMYVLLPE--NDLSEIENKLTQNLMEWTNPRMTSKYVEVFPFKIEKNYEMKQYL 291
Db 240 DLSMIVLLPNEIDGLQLEKLTAEKXLENTSLOWMRETRVLDLHPREFKVESYDLKDTL 299
QY 292 RALGLKIDFDESKADLSGASGRGLYISRMMEKSVIEVTEGTEATAAT-----GSNIVE 346
Db 300 RTMGWVDIFN-GDADLSGMTSGRLVSLGVLHKAFVEVTEGAEAAAATAVVGFGSS--- 355
QY 347 KQLPOST--LFRADHPFLFVIR--KDDILFSGKVSCP 380
Db 356 ---FTSNEEFHCNHPFLFIRQNKNTNSILFYGRFSSP 390

RESULT 14
US-10-094-944-12
; Sequence 12, Application US/10094944
; Publication No. US20030180275A1
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Clarke, Howard R
; APPLICANT: DuBoise, Robert F
; APPLICANT: Wiley, Steven R
; TITLE OF INVENTION: HUMAN SERPIN POLYPEPTIDES
; FILE REFERENCE: 3223-A
; CURRENT APPLICATION NUMBER: US/10/094,944
; CURRENT FILING DATE: 2002-03-08
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-944-12

Query Match 38.4%; Score 752; DB 12; Length 390;
Best Local Similarity 41.5%; Pred. No. 1e-60;
Matches 165; Conservative 61; Mismatches 126; Indels 26; Gaps 10;

QY 1 MASTAAANAEFCNLFREMDNDQNGNVPFSSLSLFAALALVRLGAODDSLSQIDKLHV 60
Db 1 MNSLSEANTKMPDLFOQFRKSKEN-NIFYSPISITSLGMLVLLGAKONTAQIKKVLRF 59
QY 61 -----NTASGYGNSNSQSG-LOSQKRVFSDINASHKDYDLSIVNGLFAEKVYGFHKDY 114
Db 60 DQVTEINTGKATYHVDSEGNVHHQFQKLLIEFNKSTDAVELKIANKLFGEKTYLFQY 119
QY 115 IECAEKLYDAKVERVDFTNHLIEDTRNNKXVETHGKIKNVIGEGISSAVMVLVNA 174
Db 120 LDAIKFYQTSVESVDANAPESRKKINSWESQTFNEKIKNLPEGNIGSNTTLVLVNA 179
QY 175 VYFKGWQSAFTKSTINCHFKPKSGKAVAMHQRKENLSVIEDPSMKILELRNG- 233
Db 180 IYFKGWQKXKFKNEKTEKRFWPNKNTYKSIQMMKQYTSFHFASLEDVQAKVLEIPYKX 239
QY 234 GINMYVLLPE--NDLSEIENKLTQNLMEWTNPRMTSKYVEVFPFKIEKNYEMKQYL 291
Db 240 DLSMIVLLPNEIDGLQLEKLTAEKXLENTSLOWMRETRVLDLHPREFKVESYDLKDTL 299
QY 292 RALGLKIDFDESKADLSGASGRGLYISRMMEKSVIEVTEGTEATAAT-----GSNIVE 346
Db 300 RTMGWVDIFN-GDADLSGMTSGRLVSLGVLHKAFVEVTEGAEAAAATAVVGFGSS--- 355
```

```
QY 347 KQLPOST--LFRADHPFLFVIR--KDDILFSGKVSCP 380
Db 356 ---PASTNEEFHCNHPFLFIRQNKNTNSILFYGRFSSP 390

RESULT 15
US-10-113-113-2
; Sequence 2, Application US/10113113
; Publication No. US20030166852A1
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: PROTEINASE INHIBITOR ZSERP9
; FILE REFERENCE: 00-92
; CURRENT APPLICATION NUMBER: US/10/113,113
; CURRENT FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: 60/280,678
; PRIOR FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-113-113-2

Query Match 38.3%; Score 750.5; DB 12; Length 425;
Best Local Similarity 37.4%; Pred. No. 1.e-60;
Matches 159; Conservative 86; Mismatches 135; Indels 45; Gaps 5;

QY 1 MASTAAANAEFCNLFREMDNDQNGNVPFSSLSLFAALALVRLGAODDSLSQIDKLHV 60
Db 1 MDLSVTANTKPCFDLQFEGKDDRKNIFPSLUSALGHWRLGARSDSAHQIDEVLHP 60
QY 61 NTAS-----GYGNSNSQSGSQS-QKRVF 84
Db 61 NEFSQNESKEPDCLKSNKQKVLADSSLEQKKTTPDLDDQAGSNNESGLVSCYFQQL 120
QY 85 SDINASHKDYDLSIVNGLFAEKVYGFHKDYIECAEKLYDAKVERVDFTNHLIEDTRNNK 144
Db 121 SKLDRIKTDYLTSLANRLYGEQFPICQYLDGVIOFYHTTIESVDVQKNPEKSKQEIF 180
QY 145 WVENETHGKIKNVIGEGISSAVMVLVNAVYFKGWQSAFTKSETINCHFKSPKCSGA 204
Db 181 WVECSQKIKELPSKDINAETVLVLVNAVYFAKWEYFHEHTVDAFPCLNANEKS 240
QY 205 VAMHQERKFNLSVIEDPSMKILELY-NGGINMYVLLPEND-----LSEIENKLTQNL 257
Db 241 VKWMTQKGLYRIGFIEVKAQILEMRYTKGLSMFVLLPSHSKONLKGLELEERKITHX 300
QY 258 LMEWTNPRMTSKYVEVFPFKIEKNYEMKQYLALGLKIDFDESKADLSGASGRGLY 317
Db 301 MVAMSSSENMEEESVLSFPTLEDSYDLNLSIQDMGITDIEDTRADLTGISPSNLY 360
QY 318 ISSMHKSVIEVTEGTEATAATGSIIVEKOLPOSTLFRADHPFLFVIR--KDDILFSG 375
Db 361 LSKLHKIFVVDENGIGQAAATGAVSERSLSWVEFNANHPFLFIRNKTKTLFVG 420
QY 376 KVSCP 380
Db 421 RVCSP 425

Search completed: December 12, 2003, 16:24:31
Job time : 32 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 16:20:11 ; Search time 17 Seconds
(without alignments)
1051.185 Million cell updates/sec

Title: US-09-508-997A-2

Perfect score: 1959

Sequence: 1 MASLAAANAEFCNLFREND.....FLVIRKDDILFSGKVSCTP 380

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1959	100.0	380	1 SPB7_HUMAN	Q75635 homo sapien
2	1505	76.8	380	1 SPB7_MOUSE	Q96955 mus musculu
3	819	41.8	392	1 SB11_HUMAN	Q96P15 homo sapien
4	763.5	39.0	405	1 SB12_HUMAN	Q96P63 homo sapien
5	752	38.4	390	1 SCC1_HUMAN	P93508 homo sapien
6	748	38.2	390	1 SCC2_HUMAN	P48594 homo sapien
7	739.5	37.7	379	1 ILEU_HORSE	P05619 equus cabal
8	722.5	36.9	379	1 ILEU_HUMAN	P30740 homo sapien
9	714.5	36.5	391	1 SB13_HUMAN	Q9U18 homo sapien
10	711	36.3	388	1 OVAY_CHICK	P01014 gallus gall
11	702	35.8	378	1 PT16_BOVIN	O02739 bos taurus
12	696	35.5	378	1 ILEU_FIG	P80229 sus scrofa
13	664.5	33.9	415	1 PA12_HUMAN	P05120 homo sapien
14	662	33.8	374	1 SPB8_HUMAN	P50452 homo sapien
15	658.5	33.6	397	1 SB10_HUMAN	P48595 homo sapien
16	654	33.4	385	1 OVAL_CHICK	P01012 gallus gall
17	633.5	33.4	382	1 OVAL_COTUA	P91014 coturnix co
18	649	33.1	378	1 PT16_MOUSE	Q60854 mus musculu
19	644	32.9	376	1 PT16_HUMAN	P35237 homo sapien
20	630.5	32.2	415	1 PA12_MOUSE	P12388 mus musculu
21	623	31.8	376	1 SPB9_HUMAN	P50453 homo sapien
22	622	31.8	416	1 PA12_RAT	P39524 rattus norv
23	605.5	30.9	375	1 MASP_RAT	P70584 rattus norv
24	600.5	30.7	375	1 MASP_MOUSE	P70124 mus musculu
25	592	30.2	465	1 ANT3_MOUSE	P32261 mus musculu
26	590.5	30.1	375	1 MASP_HUMAN	P36952 homo sapien
27	567	28.9	433	1 ANT3_BOVIN	P41361 bos taurus
28	566	28.9	464	1 ANT3_HUMAN	P01008 homo sapien
29	557	28.4	465	1 ANT3_SHEEP	P32282 ovis aries
30	503.5	25.7	410	1 NEUS_RAT	Q9J182 rattus norv
31	498.5	25.4	410	1 NEUS_HUMAN	Q99574 homo sapien
32	492	25.1	405	1 SPI2_MOUSE	Q9Jk88 mus musculu
33	484.5	24.7	410	1 NEUS_MOUSE	Q95684 mus musculu

```

34 478.5 24.4 417 1 KBP_MOUSE
35 476.5 24.3 410 1 NEUS_CHICK
36 476.5 24.3 423 1 AACT_HUMAN
37 467 23.8 405 1 SPI2_HUMAN
38 467 23.6 410 1 ALAT_DIDMA
39 462.5 23.6 418 1 COTR_MOUSE
40 460 23.5 418 1 SI24_APOSY
41 457.5 23.4 413 1 ALMS_TAMSI
42 455 23.2 418 1 CPI6_RAT
43 451 23.0 405 1 IPSP_MOUSE
44 446.5 22.8 405 1 ALAS_CAVPO
45 446.5 22.8 406 1 IPSP_HUMAN

```

ALIGNMENTS

```

RESULT 1
SPB7_HUMAN
ID SPB7_HUMAN STANDARD; PRT; 380 AA.
AC Q75635;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Megsin (TP55) (Serpin B7).
GN SERPINB7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97326116; PubMed=9182567;
RA Tsujimoto M., Tsuruoka N., Ishida N., Kurihara T., Iwasa F.,
RA Yamashiro K., Rogi T., Kodama S., Katsuragi N., Adachi M.,
RA Katayama T., Nakao M., Yamauchi K., Hashino J., Haruyama M., Miura K.,
RA Nakanishi T., Nakazato H., Teramura M., Mizoguchi H., Yamaguchi N.,
RA "Purification, cDNA cloning, and characterization of a new serpin with
RT megakaryocyte maturation activity."
RL J. Biol. Chem. 272:15373-15380(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Mesangial cells;
RA Miyata T., Nangaku M., Suzuki D., Inagi R., Uragami K., Sakai H.,
RA Okubo K., Kurokawa K.;
RA "A mesangium-predominant gene, megsin, is a new serpin upregulated in
RT LGA nephropathy."
RL J. Clin. Invest. 102:828-836(1998).
CC -!- FUNCTION: Might function as an inhibitor of Lys-specific
CC proteases. Might influence the maturation of megakaryocytes via
CC its action as a serpin.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- TISSUE SPECIFICITY: Predominantly expressed in mesangial cells.
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D88575; BAA31232.1; -.
CC EMBL; AF027866; AAC64506.1; -.
CC HSSP; P05619; IHLE.
CC Genew; HGNC:13902; SERPINB7.
CC MIM; 603357; -.
CC GO; GO:0004868; F:serpin; TAS.
CC InterPro; IPR000215; Serpin.
CC Pfam; PF00079; serpin; 1.
CC SMART; SM00093; SERPIN; 1.

```

```

DR PROSITE; PS00284; SERPIN; 1.
KW Serpin; Serine protease inhibitor.
FT ACT SITE 347 348 REACTIVE BOND (BY SIMILARITY).
SQ SEQUENCE 380 AA; 42904 MW; 9A2CDB6C63CFE05 CRC64;

Query Match 100.0%; Score 1959; DB 1; Length 380;
Best Local Similarity 100.0%; Pred. No. 4e-129;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASLAANAARFCNLFREMDNQNGNVFFSSLSLFAALALVRLGAGDSSQIDKLLHV 60
DB 1 MASLAANAARFCNLFREMDNQNGNVFFSSLSLFAALALVRLGAGDSSQIDKLLHV 60
QY 61 NTAGYGNSSNSQSGQLKRVFSDINASHKDYDLSVNGLFAEKYGFHGYIECAEK 120
DB 61 NTAGYGNSSNSQSGQLKRVFSDINASHKDYDLSVNGLFAEKYGFHGYIECAEK 120
QY 121 LYDAKVERVDFTHLEDTRNINKWVENETHGKIKVIVGEGGSSSAVMVLNAVYFKG 180
DB 121 LYDAKVERVDFTHLEDTRNINKWVENETHGKIKVIVGEGGSSSAVMVLNAVYFKG 180
QY 121 LYDAKVERVDFTHLEDTRNINKWVENETHGKIKVIVGEGGSSSAVMVLNAVYFKG 180
DB 121 LYDAKVERVDFTHLEDTRNINKWVENETHGKIKVIVGEGGSSSAVMVLNAVYFKG 180
QY 181 WSAFTKSETINCHFKSPKCSKAVAMHQRKFNLSVIEDPSMKILELYNGGINMYVL 240
DB 181 WSAFTKSETINCHFKSPKCSKAVAMHQRKFNLSVIEDPSMKILELYNGGINMYVL 240
QY 181 WSAFTKSETINCHFKSPKCSKAVAMHQRKFNLSVIEDPSMKILELYNGGINMYVL 240
DB 181 WSAFTKSETINCHFKSPKCSKAVAMHQRKFNLSVIEDPSMKILELYNGGINMYVL 240
QY 301 DESKADLSGIASGRILYISRMHKSYLEVTEEGTATATGNSNIVEKQLPQSTLFRADHP 360
DB 301 DESKADLSGIASGRILYISRMHKSYLEVTEEGTATATGNSNIVEKQLPQSTLFRADHP 360
QY 361 FLFVIRKDDIILFSGKVSCP 380
DB 361 FLFVIRKDDIILFSGKVSCP 380

RESULT 2
SPF7 MOUSE
ID SPF7 MOUSE STANDARD; PRT; 380 AA.
AC Q9D695;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Mesgin (Serpin B7).
GN SERPINB7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21368006; Pubmed=11473647;
RA Nangaku M., Miyata T., Suzuki D., Umezono T., Hashimoto T., Wada T.,
RA Yagi M., Nagano N., Inagi R., Kurokawa K.:
RT "Cloning of rodent mesgin revealed its up-regulation in
RT mesangio proliferative nephritis."
RL Kidney Int. 60:641-652 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STEAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=21085660; Pubmed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Caesteleand T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli P., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barch G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

```

```

Browstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
Gustincich S., Hill D., Hornam M., Hume D.A., Kamiya M., Lee N.H.,
Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
Hayashizaki Y.:
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
CC -!- FUNCTION: Might function as an inhibitor of Lys-specific
CC proteases. Might influence the maturation of megakaryocytes via
CC its action as a serpin (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF105328; AAL15768.1; -
CC EMBL; AK014524; BAB29410.1; -
CC HSSP; P05121; 1A7C.
CC MGD; MGI:2151053; Serpinb7.
CC InterPro; IPR000215; Serpin.
CC Pfam; PF00079; serpin; 1.
CC SMART; SM00093; serpin; 1.
CC PROSITE; PS00284; SERPIN; 1.
KW Serpin; Serine protease inhibitor.
FT ACT SITE 347 348 REACTIVE BOND (BY SIMILARITY).
SQ SEQUENCE 380 AA; 43050 MW; C9240272CEP9CF4 CRC64;

Query Match 76.8%; Score 1505; DB 1; Length 380;
Best Local Similarity 73.9%; Pred. No. 1.4e-97;
Matches 281; Conservative 53; Mismatches 46; Indels 0; Gaps 0;

QY 1 MASLAANAARFCNLFREMDNQNGNVFFSSLSLFAALALVRLGAGDSSQIDKLLHV 60
DB 1 MASLAANAARFCNLFREMDNQNGNVFFSSLSLFAALALVRLGAGDSSQIDKLLHV 60
QY 61 NTAGYGNSSNSQSGQLKRVFSDINASHKDYDLSVNGLFAEKYGFHGYIECAEK 120
DB 61 NTAGYGNSSNSQSGQLKRVFSDINASHKDYDLSVNGLFAEKYGFHGYIECAEK 120
QY 121 LYDAKVERVDFTHLEDTRNINKWVENETHGKIKVIVGEGGSSSAVMVLNAVYFKG 180
DB 121 LYDAKVERVDFTHLEDTRNINKWVENETHGKIKVIVGEGGSSSAVMVLNAVYFKG 180
QY 181 WSAFTKSETINCHFKSPKCSKAVAMHQRKFNLSVIEDPSMKILELYNGGINMYVL 240
DB 181 WSAFTKSETINCHFKSPKCSKAVAMHQRKFNLSVIEDPSMKILELYNGGINMYVL 240
QY 181 WSAFTKSETINCHFKSPKCSKAVAMHQRKFNLSVIEDPSMKILELYNGGINMYVL 240
DB 181 WSAFTKSETINCHFKSPKCSKAVAMHQRKFNLSVIEDPSMKILELYNGGINMYVL 240
QY 241 LPENDLSIENKLTQNLMEWTPRMTSKYVEVFPQFKIEKNYEMKQYLRAIGLKDF 300
DB 241 LPENDLSIENKLTQNLMEWTPRMTSKYVEVFPQFKIEKNYEMKQYLRAIGLKDF 300
QY 241 LPENDLSIENKLTQNLMEWTPRMTSKYVEVFPQFKIEKNYEMKQYLRAIGLKDF 300
DB 241 LPENDLSIENKLTQNLMEWTPRMTSKYVEVFPQFKIEKNYEMKQYLRAIGLKDF 300
QY 301 DESKADLSGIASGRILYISRMHKSYLEVTEEGTATATGNSNIVEKQLPQSTLFRADHP 360
DB 301 DESKADLSGIASGRILYISRMHKSYLEVTEEGTATATGNSNIVEKQLPQSTLFRADHP 360
QY 361 FLFVIRKDDIILFSGKVSCP 380
DB 361 FLFVIRKDDIILFSGKVSCP 380

RESULT 3
SB11 HUMAN
ID SB11 HUMAN STANDARD; PRT; 392 AA.
AC Q96P15; Q96P13; Q96P14;

```


DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Squamous cell carcinoma antigen 2 (SCCA-2) (Leupin).
GN SERPINE4 OR SCCA2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=95241462; PubMed=7724531;
RA Schneider S.S., Schick C., Fish K.E., Miller E., Pena J.C.,
Trer S.D., Hui S.M., Silverman G.A.;
RT "A serine proteinase inhibitor locus at 18q21.3 contains a tandem
RT duplication of the human squamous cell carcinoma antigen gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:3147-3151(1995).
RN (2)
RP SEQUENCE FROM N.A.
RX MEDLINE=96013687; PubMed=7589435;
RA Barnes R.C., Worrall D.M.;
RT "Identification of a novel human serpin gene; cloning sequencing and
RT expression of leupin.";
RL FEBS Lett. 373:61-65(1995).
RN (3)
RP SEQUENCE FROM N.A.
RX MEDLINE=21167379; PubMed=11267667;
RA Hamada K., Shinomiya H., Asano Y., Kihara T., Iwanoto M., Hanakawa Y.,
Hoshimoto K., Hirose S., Kyo S., Ito M.;
RT "Molecular cloning of human squamous cell carcinoma antigen 1 gene and
RT characterization of its promoter.";
RL Biochim. Biophys. Acta 1518:124-131(2001).
RN (4)
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schaller G.D.,
Altshul S.P., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Sapichon M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallahy S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W.,
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: MAY ACT AS A PROTEASE INHIBITOR TO MODULATE THE HOST
CC IMMUNE RESPONSE AGAINST TUMOR CELLS.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: Squamous cells.
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DE EMBL; U19576; AAA92602.1; -
DR EMBL; U19570; AAA92602.1; JOINED.
DE EMBL; U19571; AAA92602.1; JOINED.
DR EMBL; U19572; AAA92602.1; JOINED.
DE EMBL; U19574; AAA92602.1; JOINED.

DR EMBL; U19575; AAA92602.1; JOINED.
DR EMBL; U19557; AAA97553.1; -
DR EMBL; X89015; CAR61420.1; -
DR EMBL; AB035089; BAB21525.1; -
DR EMBL; BC017401; AAHL17401.1; -
DR HSSP; P01C08; IATH.
DR Genew; HGNC:10570; SERPINB4.
DR MIM; 600518; -
DR GO; GO:0005737; C:cytoplasm; NAS.
DR GO; GO:0004868; F:serpin; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR GO; GO:0030152; P:regulation of proteolysis and peptidolysis; NAS.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; serpin; 1.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
KW Serpin; Serine protease inhibitor.
FT ACT_SITE 354..355 REACTIVE BOND.
SQ SEQUENCE 390 AA; 44854 MW; 04E213CD892587D5 CRC64;

Query Match 38.2%; Score 748; DB 1; Length 390;
Best Local Similarity 40.8%; Pred. No. 6.2e-45;
Matches 160; Conservative 84; Mismatches 134; Indels 14; Gaps 8;

QY 1 MASLAANAACFVLPREMDDNGNGNVFSSLSLFAALALVRLGAQDLSQIDKLHV 60
DB 1 MNSLSFANTKFMFDFQPRKSKEN-NIFYSPISITSLGMLVGAKDNTAQQISKVLHF 59
QY 61 ----NTASGYNSSNSQSG-LQSOLKRVFSDINASHKDYDLSIVNGLFAKQVYGFHKDY 114
DB 60 DQVTNTEKATYHVDKSGNVHGFQKLLTEFNKSTDAYELKIANKLFGEKYQFLQY 119
QY 115 IECAEKLVDKVERVDFTNHEEDTRNTKWNENETHGKIKNVIGEGGISSAWMYLVNA 174
DB 120 LDAIKFQTSVESTDFANAPESKKINSWESQTNKIKLFPDGTIGNDTTLVNA 179
QY 175 VYFKGWSQATKSETINCHFSKCSOKAYAMHQERKFNLSVIEDPSMKILELYNG- 233
DB 180 IYFKGWNKFKXNTKEEFKFNKNTYKSVQMMRQVNSFNALLEDVQAKVLEIFYK 239
QY 234 GINMYVLLPE--NDLSIENKLTQNLMEWTNPRMTSKYVVEFPPOFKIKENYMKYL 291
DB 240 DLSKIVLLENDIGDQKLEKLTAKLMEWTSIQNRETCDVLHLPKFMESYDLDKTL 299
QY 292 RAUGLKDFDSKADLSGASGRLYISRMMKSYLEVTETGTEATATGTSNIVERKQLFQ 351
DB 300 RTMGWNI-FN-GDADLSGMTSGSLSVKHLKAFVETEGVEAAAAATAVWVVELSSPS 358
QY 352 ST-LREACHPTFVIR--KDDILFSGKVSCP 380
DB 359 TNEEPCNHPFFIFIRONTNLSILFYGRFSP 390

RESULT 7
ID ILEU HORSE STANDARD; PRT; 379 AA.
AC P05619;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Leukocyte elastase inhibitor (LEI).
GN SERPINE1 OR ELAH2.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=93319507; PubMed=7687128;
RA Kordula T., Dubin A., Schooltink H., Koj A., Heinrich P.C.,
RA Rose-John S.;
RT "Molecular cloning and expression of an intracellular serpin: an
RT elastase inhibitor from horse leucocytes.";

RL Biochem. J. 293:187-193(1993).
 RN [2].
 RP SEQUENCE.
 RC TISSUE=Leukocyte; PubMed=1551869;
 RX MEDLINE=92202200; PubMed=1551869;
 RA Dubin A., Travis J., Englund J.J., Potempa J.;
 RT "Equine leukocyte elastase inhibitor. Primary structure and
 identification as a thymosin-binding protein.";
 RL J. Biol. Chem. 267:6576-6583(1992).
 RN [3].
 RP SEQUENCE OF 343-362.
 RX MEDLINE=922389327; PubMed=3366785;
 RA Potempa J., Dubin A., Watorek W., Travis J.;
 RT "An elastase inhibitor from equine leukocyte cytosol belongs to the
 serpin superfamily. Further characterization and amino acid sequence
 of the reactive center.";
 RL J. Biol. Chem. 263:7364-7369(1988).
 RN [4].
 RP X-RAY CRYSTALLOGRAPHY (1.95 ANGSTROMS).
 RX MEDLINE=92389327; PubMed=1518052;
 RA Baumann U., Bode W., Huber R., Travis J., Potempa J.;
 RT "Crystal structure of cleaved equine leukocyte elastase inhibitor
 determined at 1.95-A resolution.";
 RL J. Mol. Biol. 226:1207-1218(1992).
 CC -1- FUNCTION: THIS INHIBITOR IS THOUGHT TO BE INVOLVED IN THE CONTROL
 OF INTRACELLULAR PROTEIN TURNOVER. BINDS TO THYMOSIN BETA-4.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 DR ENBL; M91161; AAA97513.1; -.
 DR PIR; A28060; A28060.
 DR PIR; A2421; A2421.
 DR PDB; 1HLE; 31-JAN-94.
 DR InterPro; IPR000215; Serpin.
 DR Pfam; PF00079; serpin; 1.
 DR SMART; SM00093; SERPIN; 1.
 DR PROSITE; PS00284; SERPIN; 1.
 KW Serpin; Serine protease inhibitor; 3D-structure.
 FT MOD_RES 1 1
 FT ACT_SITE 344 345
 FT HELIX 2 22
 FT STRAND 28 30
 FT HELIX 32 44
 FT TURN 45 45
 FT HELIX 48 58
 FT TURN 59 59
 FT HELIX 60 62
 FT TURN 64 65
 FT HELIX 66 77
 FT TURN 78 78
 FT STRAND 85 95
 FT TURN 96 97
 FT HELIX 102 112
 FT STRAND 115 119
 FT TURN 121 123
 FT HELIX 125 139
 FT TURN 140 142
 FT TURN 150 151
 FT TURN 155 156
 FT STRAND 157 171
 FT HELIX 177 179
 FT STRAND 181 186
 FT TURN 192 209
 FT STRAND 210 212
 FT HELIX 210 212
 FT TURN 213 213

FT STRAND 214 221
 FT TURN 222 223
 FT STRAND 226 233
 FT HELIX 244 248
 FT TURN 249 249
 FT HELIX 252 259
 FT HELIX 261 263
 FT STRAND 265 274
 FT STRAND 276 283
 FT HELIX 285 291
 FT TURN 292 292
 FT HELIX 295 297
 FT TURN 299 301
 FT HELIX 305 308
 FT STRAND 313 326
 FT STRAND 330 343
 FT STRAND 350 353
 FT STRAND 358 364
 FT TURN 365 368
 FT STRAND 369 376
 SQ SEQUENCE 379 AA; 43006 MW; 72DED1999371427A CRC64;
 Query Match 37.7%; Score 739.5; DB 1; Length 379;
 Best Local Similarity 38.0%; Pred. No. 2.3e-44;
 Matches 146; Conservative 85; Mismatches 137; Indels 19; Gaps 4;
 QY 1 MASLAANAEEFCNLFREMDNQNGNVFFSSLSLFAALVRLGAQDLSQIDKLHV 60
 DB 1 MEQLSTANTHFAVDFRALNESDPTGNIPISPLSISSALAMIFLGRNTAAQVSKALYF 60
 QY 61 NTASGYGNSNSQSGLSQKLVPSDINASHKDYDLSIVNGLFAEKVYGFHKDYIECAEK 120
 DB 61 DTV-----EDIHSPQSLNADINKPGAPYIIKLANLYGKTYNFIADFLASTOK 110
 QY 121 LYDAKVERVDFTNLETRNINKWVENETHGKIKNVI GEGSISSAMVVLVNAVYFKGK 180
 DB 111 MYGASLAVDQQAPEDARKEINENWKGTEGKIPELLVKGWVDNMTKLVNAVYFKGN 170
 QY 181 WQSAFTKSETINCHFKSPKCSGKAVAMQHQRKPNLSVIEDSKMKILELYNG-GINMYV 239
 DB 171 WQEKPMKEATRDAPFLNKDKDTKTVKMYQKKFPYNYIEDLKRVLELPYQKELSMII 230
 QY 240 LLPEN-----DLSEIENKLTQNLMEWTNPRMTSKYVEVFFQFKIKNYEMKQYLRA 293
 DB 231 LLPDDIDESTGLSKIEKQTLLEKREWKPNLYLAENVNHLPRKLESYDLTSLHAR 290
 QY 294 LGLKDIFDSKADLSGASGRLYISRMMHKSIVETEGTEATAATGSGNIVKQLPQST 353
 DB 291 LGVQDLFNRGRADLSGMSGARDLFVSKIIHKSPVDLNEGTEAAAAATAGTITLAWLMPPE 350
 QY 354 LFRADHPFLFVIRKDDI--ILFSGKVSCP 380
 DB 351 NFNADHPFIFFIRNPSANILFLGRFSSP 379
 RESULT 8
 ID ILEU HUMAN STANDARD; PRT; 379 AA.
 AC P30740;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE leukocyte elastase inhibitor (LEI) (Monocyte/neutrophil elastase
 inhibitor) (W/NEI) (BI).
 GN SERPINE1 OR ELANH2 OR PI2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92302296; PubMed=1376927;
 RA Remold-O'Donnell E., Chin J., Alberts M.;


```

RI Biochim. Biophys. Acta 1492:441-446 (2000).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Placenta;
RX MEDLINE=21213570; PubMed=11313015;
RA Acts H.F., Weiss T., Scheuring S., Scott F.J., Irving J.A., Michel G.,
RD Bird P.L., Ruzicka T.;
RE "Sequence, organization, chromosomal localization and alternative
RT splicing of the human serine protease inhibitor gene hurpin (P113),
RT which is up-regulated in psoriasis.";
RL DNA Cell Biol. 20:123-131 (2001).
CC -|- FUNCTION: May play a role in the proliferation or differentiation
CC of keratinocytes.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC -|- ALTERNATIVE PRODUCTS:
CC Evert=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q9UIV8-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9UIV8-2; Sequence=VSP_006058;
CC TISSUE SPECIFICITY: SKIN-SPECIFIC.
CC -|- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ001696; CAA04935.2; -;
DR EMBL; AJ001697; CAA04936.2; -;
DR EMBL; AJ001698; CAA04937.1; -;
DR EMBL; AF169949; AAD55765.1; -;
DR EMBL; AF216854; AAF72879.1; -;
DR EMBL; AJ278717; CAC03569.1; -;
DR PIR; JC7118; JC7118.
DR HSSP; P05120; 1BY7.
DR Genew; HGNC:8944; SERPINB13.
DR MIM; 604445; -;
DR GO; GO:0004868; F:serpin; NAS.
DR GO; GO:0030162; P:regulation of proteolysis and peptidolysis; NAS.
DR GO; GO:0009411; P:response to UV; TAS.
DR InterPro; IPR000240; Maspin.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; serpin; 1.
DR PRINTS; PR00676; MASPIN.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
KW Serpin; Serine protease inhibitor; Alternative splicing.
FT ACT SITE 356 357 REACTIVE BOND (BY SIMILARITY).
FT VARSPIC 206 257 Missing (in isoform 2).
FT FTID=VSP_006058.
FT CONFLICT 8 8 S -> N (IN REF. 4).
FT CONFLICT 75 75 MISSING (IN REF. 1; CAA04937).
FT CONFLICT 293 293 G -> S (IN REF. 2).
FT CONFLICT 297 297 E -> Q (IN REF. 1; CAA04937).
FT SEQUENCE 391 AA; 44276 MW; 2CA88558D4BC2B09 CRC64;
SQ
Query Match 36.5%; Score 714.5; DB 1; Length 391;
Best Local Similarity 38.0%; Pred. No. 1.3e-42;
Matches 152; Conservative 77; Mismatches 142; Indels 29; Gaps 7;
QY 1 MASIAAANAFCNLFREMDNDGNGNFFSLSLFAALVALVRLGAQCSLSQIDKLLHV 60
DB 1 YDSIGAVSTRICFLDFELXEL-KKINDGNIFPSPVGLITAGVLLTGATASQLEVEFH- 58
QY 61 NTASGVGNSNSQSGLSQKRVFSDINASHK-----DYDLSLVNGLPAPK 106
DB 59 -----SEKETKSRIKAEKEKVIENIEAVHQCPQKFLTEISKLTNDYELNITNRLFGK 112
QY 107 VYGFHKDYECABKLYDAKVERVDYFNHLEDTRETNINKWENETGKIKNVIGEGGISS 166

```

```

DB 113 TVFLQKLYDYVEKYHSLPVDVNADESRKKNWSKTKNKLKDLFPDGISS 172
QY 167 AMVLVNAVYFKGQWOSAPFTKSETINCHFKSPKCSGKAVAMHQERKFNLSVIEDPSMKI 226
DB 173 TKLVVNAVYFKQWDBREKFKENTKEEFWANKSTSKSVQMTQSHSFTEFLDLQAKI 232
QY 227 LELV-NGGNNVYLLPNDLSEIE---NKLTFQNLMENTNPRMTSKVVEVFPQKIE 282
DB 233 LGIPYXNDLSMVELLP-NDIDGLEKIIDKISPEKLVWNTSPGHMEERKVNHLRPFVE 291
QY 283 KXVENKQVLRALGLKIDPESKADLSGAGRLYISMMHKSIVIEVTEGTEAATGS 342
DB 292 DGYDLEAVLAMGNGDAPSEKADYSGMSSGGLYAKFLHSSFVATBEGTEAATAAGI 351
QY 343 NIVEKQLPQSTLPRADHPFLFVIR--KDDILLFSKGVSCP 380
DB 352 GFTVTSAPGHENVECHNPFPLFIRHNSNISILFFGRFSSP 391
RESULT 10
OVAY CHICK
ID OVAY CHICK STANDARD; PRT; 388 AA.
AC P01014;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE Gene Y protein (Ovalbumin-related).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83014329; PubMed=7122240;
RA Heilig R., Muraskovsky R., Kloepper C., Mandel J.L.;
RT "The ovalbumin gene family: complete sequence and structure of the Y
RT gene.";
RL Nucleic Acids Res. 10:4363-4382 (1982).
CC -|- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J00922; AAA68882.1; -;
DR PIR; A01244; DYCH.
DR HSSP; P01012; LOVA.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; serpin; 1.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
KW Serpin.
FT ACT SITE 353 354 REACTIVE BOND HOMOLOG.
FT CARBOHYD 95 95 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 215 215 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 312 312 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 388 AA; 43772 MW; 2AF17BF715D7F461 CRC64;
Query Match 36.3%; Score 711; DB 1; Length 388;
Best Local Similarity 39.6%; Pred. No. 2.3e-42;
Matches 155; Conservative 75; Mismatches 147; Indels 14; Gaps 7;
QY 1 MASIAAANAFCNLFREMDNDGNGNFFSLSLFAALVALVRLGAQCSLSQIDKLLHV 60
DB 1 MDSISVNAKFCDFVFMKVVHVNILYCPLSILFALAMVYLARGNTESQMKVLHF 60

```

[illegible]

```

RESULT 11
PT16_BOVIN
ID   FTT16_BOVIN          STANDARD;          PRT;   378 AA.
AC   C02739;
DT   30-MAY-2000 (Rel. 39, Created)
DI   30-MAY-2000 (Rel. 39, Last sequence update)
DR   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Serine proteinase inhibitor B-43.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Rutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovidae; Bovinae; Bos.
NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RT   TISSUE=Brain;
RX   MEDLINE=97165893; PubMed=9013786;
RA   Nakaya N., Nishibori M., Kawabata M., Saeki K.;
RT   "Cloning of a serine proteinase inhibitor from bovine brain:
RT   expression in the brain and characterization of its target
RT   proteinases";
RL   Brain Res. Mol. Brain Res. 42:293-300 (1996).
CC   -!- FUNCTION: MAY BE INVOLVED IN THE REGULATION OF SERINE PROTEINASES
CC   PRESENT IN THE BRAIN OR EXTRAVASATED FROM THE BLOOD.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- TISSUE SPECIFICITY: BRAIN.
CC   -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
CC   EMBL; D55670; BAAL19875.1; -.
CC   HSPG; P05120; 1BY7.
DR   InterPro; IPR000215; Serpin.
DR   Pfam; PF00079; serpin; 1.
DR   SMART; SM00093; SERPIN; 1.
DR   PROSITE; PS00284; SERPIN; 1.
KW   Serpin; Serine protease inhibitor.
FT   ACT_SITE 343..344 REACTIVE BOND [BY SIMILARITY].
SQ   SEQUENCE 378 AA; 42560 MW; 664F499CFCFE263A CRC64;
Query March 35.8%; Score 702; DB 1; Length 378;

```

Best Local Similarity	39.4%;	Pred. No. 9.4e-42;	
Matches 152;	Conservative	83;	Mismatches 137; Indels 14; Gaps 7;
Qy	1	MASIAAANAECFNLFREMDNQNG--NVFFSGSLFAALVRLGAQDDLSIQDKLH	59
Db	1	MDALSFANGTALTLKKL--GEGNSKWFIPLSISALAMVLGAKGNTAAQMCQTLS	58
Qy	60	VNTAGYGNSSNGSQGLQKRVFSDINASHKDYDLSIVNGLFAKVVYGHKDYIECAE	119
Db	59	LKNSGGG--RDVHQFQNL---SEVARRDTQVLLRTANRLFGKTYDPLSGFKDSCH	112
Qy	120	KLYDAKVRDVFTHIEDFRRNINKWVENETHGKIKNVIHGEGISSSAVMYLNNAVYK	179
Db	113	KFTQRMGELDFSAEQGRKHINTWAETGKIRDLFLFANSVPMTRVLVNAVYFKG	172
Qy	180	KWQSAFTKSETINCHFKPCKSGKAVAMQHOERKNLSVIEDPSMKILLELYNG--GINMY	238
Db	173	NWDTQFNKEHTESRPFVSKNVKPEVQVMFKKSTCKTYIGBISTQILVLPYVQBELNV	232
Qy	239	VLLP--ENDLSBENKLTQNLMEWTNPRMTSKTVYVFPQFKIEKVENYKQVYLAQLG	296
Db	233	ILPSESIDLNTVEKALTVEKFIAMTPDWNDEEVEVFLPRTTEESYDMEEFQELGM	292
Qy	297	KDIFDESADLSGTSAGGLYITSMWHSKSVIEVTEGGTEATATGNSVTEKQLPQSTLFR	356
Db	293	TDAFETRADPSGMSGRGLHLSKVMKHSFVEVTEGGTEAAAATGAVVMCLMVTVPFN	352
Qy	357	ADHPFLFVIR--XDDILFSGKVSQCP	380
Db	353	ANEHPFLFIQHSKTGALLFCGRFCSP	378
RESULT 12			
ILEU	PIG		
AD	ILEU	PIG	STANDARD; PRT; 378 AA.
IC	PA0229;		
DT	01-OCT-1993	(Rel. 27, Created)	
DT	01-OCT-1993	(Rel. 27, Last sequence update)	
DT	16-OCT-2001	(Rel. 40, Last annotation update)	
DE	Leucocyte elastase inhibitor (LEI)	(Leucocyte neutral protease inhibitor) (LNPI).	
DS	inhibitor	(LNPI).	
GS	SERPINE1 OR ELANE2.		
OS	Sus scrofa (Pig)		
OC	Zukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.		
OC	NCBI TaxID=9823;		
RP	[1]		
RF	SEQUENCE.		
RC	TISSUE=Leukocyte;		
FX	MEDLINE=94039065; PubMed=7901009;		
RA	Teschauer W.F., Mentele R., Sommerhoff C.P.;		
RL	"Primary structure of a porcine leukocyte serpin."		
RT	Eur. J. Biochem. 217:519-526(1993).		
CC	-!- FUNCTION: REGULATES THE ACTIVITY OF THE NEUTROPHIL PROTEASE		
CC	ELASTASE, CATHEPSIN G AND PROTEINASE-3 (By similarity).		
CC	-!- SUBUNIT: Homodimer; disulfide-linked.		
CC	-!- SUBCELLULAR LOCATION: Cytoplasmic.		
CC	-!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.		
DR	HSP; P05619; 1HL5.		
DR	InterPro; IPR000215; Serpin.		
DR	Pfam; PF00079; serpin; 1.		
DR	SMART; SM00093; SERPIN; 1.		
DR	PROSITE; PS00284; SERPIN; 1.		
FT	Serpin; Serine protease inhibitor.		
KW	MOD_RES	1	BLOCKED.
FT	FT	1	
FT	DISULFID	80	INTERCHAIN (PROBABLE).
FT	ACT_SITE	343 344	REACTIVE BOND (BY SIMILARITY).
SQ	SEQUENCE	378 AA;	42512 MW; 3CA7C65F11E508CA CEC64;

Query Match 35.5%; Score 696; DB 1; Length 378;
 Best Local Similarity 37.5%; Pred.No. 2.4e-41;
 Matches 146; Conservative 82; Mismatches 141; Indels 20; Gaps 5;

QY 1 MASLAANAFCNLPREMDDNQCNGNVFFSLSLFAALALVRLGAQDDLSQIDKILHV 60
DB 1 MEQLSAANTRFALDIFRALNESNPAGNIFISPGISSALAMILLGTRNGTEAQMVKALHF 60
QY 61 NTASGVCNSENQSLQKRVFSDINASHKDYDLSIVNGLFAEKVYGHKOYIECAEK 120
DB 61 DTV-----KDIHSRFSQSLNADINCKGASVILKLANLFGKERTHFFLPEFLASTQK 110
QY 121 LYDAKVERDFTNHLIEDFRNINKWVENETHGKIKNVIGRGITSSSAVWVIAVNAVFKGK 180
DB 111 TYGALASVDPLASEEARKALNEWKEQTEGKPELLASGVDSATKVLVNAIVFKGS 170
QY 181 WQSFTKSTTINCHFKSKGKAVAMFQEKENLNSVIEDPSMKILRLNNG-GINCVY 239
DB 171 WOEFMTFATKDAPELNLKDKSKTVKMYQKKKPPFGYIKELKORVELPQGGKDLSMVI 230
QY 240 LIPEN-----DLSIEINKLTFQNLMEWNPRTMTSKVVEFPQFKTEKQVEMKQVIRA 293
DB 231 LIPDSIEDSTGKRLKIOHUTLEKLNWTKPDNLELVNHLPRFLRESYDLNAPLAR 290
QY 294 LGLKDIFDESADLSGASGRLYISRMVHKSYIEVTEGTEATAATGTSNIVEKQLPQST 353
DB 291 LGVQDLFG-SRADLTGSGEARDLFISKVHKAFVNEEGTEAAATAGIAVFAMLMPEE 349
QY 334 LFRADHPFLVTRKD--DILFSGKVSCEP 380
DB 350 DFIADHPFIPIRHNPSNIFLGLSSP 378
RESULT 13
PAI2 HUMAN STANDARD; PRT; 415 AA.
ID AC P05120; Q96E36;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Plasmimogen activator inhibitor-2 precursor (PAI-2) (Placental
DE plasmimogen activator inhibitor) (Monocyte Arg-serpin) (Urokinase
DE inhibitor).
GN SERPINE2 OR PAI2 OR PLANH2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=67137674; PubMed=3029122;
RA Ye R.D., Wan T.-Z., Sadler J.E.,
RT "cDNA cloning and expression in Escherichia coli of a plasminogen
RT activator inhibitor from human placenta.";
RL J. Biol. Chem. 262:3718-3725 (1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89142852; PubMed=3325928;
RA Schleuning W.-D., Medcalf R.L., Hession C., Rothenbuhler R.,
RA Shaw A., Kruthof E.K.O.,
RT "Plasminogen activator inhibitor 2: regulation of gene transcription
RT during phorbol ester-mediated differentiation of U-937 human
RT histiocytic lymphoma cells.";
RL Mol. Cell. Biol. 7:4564-4567 (1987).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=89174589; PubMed=2494165;
RA Ye R.D., Aherm S.M., le Beau M.M., Lebo R.V., Sadler J.E.;
RT "Structure of the gene for human plasminogen activator inhibitor-2.
RT The nearest mammalian homologue of chicken ovalbumin.";
RL J. Biol. Chem. 264:5495-5502 (1989).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Monocytes;
RX MEDLINE=87252928; PubMed=3496414;
RA Webb A.C., Collins K.L., Snyder S.F., Alexander S.J.,
Rosenwasser L.J., Eddy R.L., Shows T.B., Auron P.E.;
RT "Human monocyte Arg-Serpin cDNA. Sequence, chromosomal assignment,
RT and homology to plasminogen activator-inhibitor.";
RL J. Exp. Med. 166:77-94 (1987).
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE=Monocytes;
RX MEDLINE=88125032; PubMed=3257578;
RA Antalis T.N., Clark M.A., Barnes T., Lehrbach P.R., Devine P.L.,
RA Chezhov G., Goss N.H., Stephens R.W., Tolstoshev P.;
RT "Cloning and expression of a cDNA coding for a human monocyte-derived
RT plasminogen activator inhibitor.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:985-989 (1988).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=90152678; PubMed=2303256;
RA Samia J.A., Alexander S.J., Horton K.W., Auron P.E., Byers M.G.,
RA Shows T.B. Jr., Webb A.C.;
RT "Chromosomal organization and localization of the human urokinase
RT inhibitor gene: perfect structural conservation with ovalbumin.";
RL Genomics 6:159-167 (1990).
RN [7]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Brance C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Besak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Pahay J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bonfield G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [8]
RP SEQUENCE OF 12-17; 103-108 AND 314-321.
RX MEDLINE=93162043; PubMed=1286667;
RA Rasmussen H.H., van Damme J., Puype M., Gesser B., Celis J.E.,
RA Vandekerckhove J.;
RT "Microsequences of 145 proteins recorded in the two-dimensional gel
RT protein database of normal human epidermal keratinocytes.";
RL Electrophoresis 13:960-969 (1992).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=99148112; PubMed=10368272;
RA Harrop S.J., Jankova J., Coles M., Jardine D., Whittaker J.S.,
RA Gould A.R., Meister A., King G.C., Mabbott B.C., Curmi P.M.G.;
RT "The crystal structure of plasminogen activator inhibitor 2 at 2.0-A
RT resolution: implications for serpin function.";
RL Structure 7:43-54 (1999).
RN [10]
RP VARIANTS ASP-120; HYS-229; LYS-404 AND CYS-413.
RX MEDLINE=99318093; PubMed=1039209;
RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemes J., Ziaugra L.,
RA Friedland L., Rolfe A., Warrington J., Lipschutz R., Daley G.Q.,
RA Lander E.S.;
RT "Characterization of single-nucleotide polymorphisms in coding regions
RT of human genes.";
RL Nat. Genet. 22:231-238 (1999).
RN [11]
RP ERRATUM.

RA Cargill M., Altmuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 RA Shaw N., Lane C.R., Lim B.P., Kalyanaram N., Nimesh J., Ziaugra L.,
 RA Friedland L., Rolfe A., Warrington J., Lipschutz R., Daley G.O.,
 RA Lander E.S.;
 RL Nat. Genet. 23:373-373(1999).
 CC -!- FUNCTION: PAI-2 INHIBITS UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
 CC THE MONOCYTE DERIVED PAI-2 IS DISTINCT FROM THE ENDOTHELIAL
 CC CELL-DERIVED PAI-1.
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC OR EXTRACELLULAR.
 CC -!- PTM: THE SIGNAL SEQUENCE IS NOT CLEAVED.
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; J02685; AAA36413.1; -;
 CC EMBL; M18082; AAA60006.1; -;
 CC EMBL; Y00630; CAA68666.1; -;
 CC EMBL; M24657; AAB60348.1; -;
 CC EMBL; M24651; AAB60348.1; JOINED.
 CC EMBL; M24652; AAB60348.1; JOINED.
 CC EMBL; M24653; AAB60348.1; JOINED.
 CC EMBL; M24654; AAB60348.1; JOINED.
 CC EMBL; M24655; AAB60348.1; JOINED.
 CC EMBL; M24656; AAB60348.1; JOINED.
 CC EMBL; J03603; AAB60004.1; -;
 CC EMBL; M31551; AAA36797.1; -;
 CC EMBL; M31547; AAA36797.1; JOINED.
 CC EMBL; M31548; AAA36797.1; JOINED.
 CC EMBL; M31549; AAA36797.1; JOINED.
 CC EMBL; M31550; AAA36797.1; JOINED.
 CC EMBL; B0012609; AAB12609.1; -;
 CC EMBL; A21238; CAA01535.1; -;
 CC EMBL; A21254; CAA01539.1; -;
 CC PIR; A32853; A32853.
 CC PDB; 1BY7; 24-OCT-99.
 CC PDB; 1JRR; 18-DEC-02.
 CC Aarhus/Shent-2DPAGE; 6314; IEF.
 CC Genew; HGNC:8584; SERPIN2.
 CC MIM; 173390; -;
 CC GO; GO:0006916; P:anti-apoptosis; TAS.
 CC InterPro; IPR000215; Serpin.
 CC Pfam; PF00079; serpin; 1.
 CC SMART; SM00093; SERPIN; 1.
 CC PROSITE; PS00284; SERPIN; 1.
 KW Serpin; Serine protease inhibitor; Plasma; Plasminogen activation;
 KW Glycoprotein; Signal; 3D-structure; Polymorphism.
 FT SIGNAL 1 ?
 FT CARBOHYD 75 75 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 115 115 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT ACT_SITE 380 381 REACTIVE BOND.
 FT VARIANT 120 120 N -> D (IN dbSNP:6098).
 FT VARIANT 229 229 R -> H (IN dbSNP:6100).
 FT VARIANT 404 404 N -> K (IN dbSNP:6103).
 FT VARIANT 413 413 S -> C (IN dbSNP:6104).
 FT VARIANT 170 170 /FTIG=VAR_011745.
 FT VARIANT 4 22 N -> Y (IN REF. 7).
 FT STRAND 28 30
 FT HELIX 32 44
 FT TURN 45 45
 FT HELIX 48 57
 FT TURN 58 63
 FT HELIX 100 102

FT HELIX 103 114
 FT STRAND 123 132
 FT TURN 133 134
 FT STRAND 137 137
 FT HELIX 139 149
 FT STRAND 154 156
 FT HELIX 158 176
 FT TURN 177 179
 FT TURN 182 183
 FT TURN 187 188
 FT TURN 192 193
 FT STRAND 196 204
 Query Match 33.9%; Score 664.5; DB 1; Length 415;
 Best Local Similarity 34.7%; Pred. No. 4.3e-39;
 Matches 144; Conservative 83; Mismatches 153; Indels 35; Gaps 5;
 QY 1 MASLAAANAECFNLPREDDNQGNGNVFSSLSLPAALVRLGAQDLSLQDKLHV 60
 DB 1 MEDLCVANTLFLALNLFKHLAKASPTQNLFLSPWSISSTWAMYMGSRGSTDQMAKVLQF 60
 QY 61 N-----TASGY-----CN-----SSNSQSLOSQKRVFSINASHKD 93
 DB 61 NEVCANAVTMTPTENFTSCGFMQIQKQSYDAILQAQADKIHSFSSLSAINASTGN 120
 QY 94 YDLSVINGLFAEKYGFHKDYIECAEKLYDAKVERVDFTNHLEDTRNNKKNVENETHGK 153
 DB 121 YLLESVNLFGKSGASPREYIRLQKYSEPOAVDFLECAEARKKINSVKTQTKGK 180
 QY 154 IKNVIGEGISSAVMLNNAVYFKKQSAFTKSTINCHFKSPKSGKAVAMHEQERK 213
 DB 181 IPNLPESSVDGDRWLVNNAVYFKGKWKTEFKCNGLYPFRVNSAQRTFVQWYLREK 240
 QY 214 ENLSVIEDPSMKILELYNGGINMYVLPE-----NDLSEIKLTFQNLMEWNTPRM 267
 DB 241 LNLGYIEOLKAQILELPPYAGDYSMFLPLPDLADVSTGLLESEITYDKLNTKTSKDN 300
 QY 268 TSKYVEVFPQFKIEKNYEMKQYLRALGLKIDFDESKADLSGIASGGRLXISRMHKSXI 327
 DB 301 AEDEVVYIPQKLEEHYELRSLASGMEDAFNKGANFSGMSERNDLFLSEVFFHQAMV 360
 QY 328 EYTEGTATATGTSNIVEKQLPQSLPRADHPELVI--RKDDILLFSGKVSQP 380
 DB 361 DVNEGTAAAGTGGVMTGTGHHGPGQFVADHPFLIMHKITNCILIFFGFGFSP 415
 RESULT 14
 SPB8 HUMAN
 ID SPB8 HUMAN STANDARD; PRT; 374 AA.
 AC P50452;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cytoplasmic antiprotease 2 (CAP2) (CAP-2) (Protease inhibitor 8)
 DE (Serpin B8).
 GN SERPINA8 OR P18.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=96102039; PubMed=8530382;
 RA Sprecher C.A., Morgenstern K.A., Mathewes S., Dahlen J.R.,
 RA Schrader S.K., Foster D.C., Kiesel W.;
 RT "Molecular cloning, expression, and partial characterization of two
 RT novel members of the ovalbumin family of serine proteinase
 RT inhibitors.";
 RL J. Biol. Chem. 270:29854-29861(1995).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: L40377; AAC41939.1; --
 DR PIR: A59273; A59273.
 DR HSP: P05120; IRY7.
 DR Genew; HGNC:8952; SERPINB8.
 DR MIM: 601697; --
 DR GO: GO:0005829; C:cytosol; TAS.
 DR GO: GO:0005515; F:protein binding activity; TAS.
 DR GO: GO:0004868; F:serpin; TAS.
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00079; serpin; 1.
 DR SMART; SM00093; SERPIN; 1.
 DR PROSITE; PS00284; SERPIN; 1.
 KW Serpin; Serine protease inhibitor.
 FT ACT_SITE 339 340 REACTIVE BOND (BY SIMILARITY).
 SQ SEQUENCE 374 AA; 42786 MW; E855A033519AC60E CRC64;

Query Match 33.8%; Score 662; DB 1; Length 374;
 Best Local Similarity 37.4%; Pred. No. 5.6e-39;
 Matches 145; Conservative 79; Mismatches 142; Indels 22; Gaps 8;

QY 1 MASLAANAAPFCNLFREMDNQGNGVFPSSLSLFAALALVPLGAGDLSQIDKLHV 60
 DB 1 MDDLCRANGTFAISLFGILGEEDNSRVFPMSISSALANVPFGKAGSTAAQWSQALCL 60
 QY 61 NTASGYGNSNSGSLQKRVFSDINASHKDYDLSTVNGLFAEKVYGFHKVIECAEK 120
 DB 61 Y-----KGDTHRGFSLL-----SEVNRGTGYQLLRTANKLFGEXTCDPLDFEYCYQK 110
 QY 121 LYDAKVRVDFTNHLDETRNKNWENETHGKIKNVIGRGGISSANWLNNAVYFKG 180
 DB 111 FYQAELELFAEDTECRKHINDWAEKTEGKI SEVIDAGTVDP LTKLVNNAVYFKG 170
 QY 181 WSAFTKSETINCHFKSPKCGKAVAMHQRKFNLSVIEDPDKMLLELY-NGGINVY 239
 DB 171 WNEQFDKRYTRGMLFKTNE-EKKVYQMFKEAFKNGYADEVITQVLELPYVESELSWI 229
 QY 240 LLPEN--DLSEIENKLTQNLMEWTNPRMTSKYVEVFPQFKTEKYEMKQYLALGLK 297
 DB 230 LLPDNDTDLAVVEKALTYEKKAWNSEKLTGSKVQVFLPLKLESDYDLEPFLRLGMI 289
 QY 298 DIFDESADLSGASGRLYISRMHKSIVIEVTEGTEATAATG---SNIVEKQLPOSTL 354
 DB 290 DAFDEAKADFGSMSTKXNPLSKVAHKCFVEVNEEGTEAANAATVVRNRCRMEPR--- 346
 QY 355 FRADHPFLFVIR--KDDILFSGKVSCP 380
 DB 347 FADHPFLFIRHKNTNCLFCGRFSP 374

RESULT 15

SE10 HUMAN
 ID SE10 HUMAN STANDARD; PRT; 397 AA.
 AC P48595;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Bonapin (Protease inhibitor 10) (Serpin B10).
 GN SERPINB10 OR P110.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN NCBI_TaxID=9606;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Bone marrow;

RX MEDLINE=96070759; PubMed=7592909;
 RA Riewald M., Schlegel R.R.;
 RT "Molecular cloning of bonapin (protease inhibitor 10), a novel human
 RT serpin that is expressed specifically in the bone marrow.";
 RL J. Biol. Chem. 270:26754-26757(1995).
 CC -1- FUNCTION: MAY PLAY A ROLE IN THE REGULATION OF PROTEASE
 CC ACTIVITIES DURING HEMATOPOIESIS.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- TISSUE SPECIFICITY: EXPRESSED SPECIFICALLY IN THE BONE MARROW.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: U35459; AAC50282.1; --
 DR PIR: I39184; I39184.
 DR HSP: P05619; IHLE.
 DR Genew; HGNC:8942; SERPINB10.
 DR MIM: 602058; --
 DR GO: GO:0005515; F:protein binding activity; TAS.
 DR GO: GO:0004868; F:serpin; TAS.
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00079; serpin; 1.
 DR SMART; SM00093; SERPIN; 1.
 DR PROSITE; PS00284; SERPIN; 1.
 KW Serpin; Serine protease inhibitor.
 FT ACT_SITE 362 363 REACTIVE BOND (BY SIMILARITY).
 SQ SEQUENCE 397 AA; 45402 MW; 8CE01246867154DF CRC64;

Query Match 33.6%; Score 658.5; DB 1; Length 397;
 Best Local Similarity 34.9%; Pred. No. 1.1e-38;
 Matches 139; Conservative 89; Mismatches 151; Indels 19; Gaps 5;

QY 1 MASLAANAAPFCNLFREMDNQGNGVFPSSLSLFAALALVPLGAGDLSQIDKLHV 60
 DB 1 MDSLATSINQFALESKSLASAGQKNIFPSSWSISTSLTIVLGAKGTAAQMAQVQLQF 60
 QY 61 NTASGYG-----NSSNSQSLQSKRVFSDINASHKDYDLSTVNGLFAEKV 107
 DB 61 NDDQVKCDPESKKKKEKFNLSSEI-IHSDPQLI SEILKPNDDVLLKTNALYGEKT 119
 QY 108 YGFHKDYIECAEKLYDAKVRVDFTNHLDETRNKNWENETHGKIKNVIGEGISSA 167
 DB 120 YAFHNKYLEDMKTVFGAEPQPVNFVEASDQIRKXINSWVERQTEGKIQNLIPDSVDSTT 179
 QY 168 VMVLNAVYFKGWQSAFTKSETINCHFKSPKCGKAVAMHQRKFNLSVIEDPDKML 227
 DB 180 RMILNAVLYFKGIWEHQFLVQNTTEKPFRTNETTSKPVQMMFKKLIHPIHEKPKAVGL 239
 QY 228 ELRYNG-GINNVILLPE--NDLSEIENKLTQNLMEWTNPRMTSKYVEVFPQFKIKN 284
 DB 240 QLYYKSRDLSSLILLPEDINGLEQLKALTYEKNLWNTSADMWELYEVLHLPKFKLSDS 299
 QY 285 YEMKQYLALGLKIDFDESKADLSGASGRLYISRMHKSIVIEVTEGTEATAATGSI 344
 DB 300 YDLKSTLSSMGMSDAFSQSKADFGSGMSARNFLNLSVNFHKAFAVEINQGTGAAGSGSEI 359
 QY 345 VEKQPLPQSTLFRADHPFLFVIR--KDDILFSGKVSCP 380
 DB 360 DIRIRVPSIEFNANHPFLFIRHKNTNCLFYGRLCSP 397

Search completed: December 12, 2003, 16:25:58

Job time : 19 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 16:23:17 ; Search time 41 Seconds
(without alignments)
2391.707 Million cell updates/sec

Title: US-09-508-997A-2

Perfect score: 1959
Sequence: 1 MASLAANAAPCNLFREMD.....FLVIRKDDIILFSGKVSQSCP 380

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL 23.*

- 1: sp_archea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rviro.*
- 16: sp_bacteriap.*
- 17: sp_archeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1505	76.8	380	11 Q9D695	Q9D695 mus musculus
2	1488	76.0	380	11 Q920J5	Q920J5 rattus norv
3	800	40.8	388	11 Q91212	Q91212 mus musculus
4	798	40.7	388	11 Q9C0V3	Q9C0V3 mus musculus
5	763.5	39.0	423	11 Q9D7P9	Q9D7P9 mus musculus
6	734	37.5	390	4 Q81X13	Q81X13 homo sapien
7	710.5	36.3	379	11 Q8BK60	Q8BK60 mus musculus
8	708.5	36.2	379	11 Q9D154	Q9D154 mus musculus
9	707	36.1	382	11 Q8VHP7	Q8VHP7 mus musculus
10	700.5	35.8	379	11 Q9D7S8	Q9D7S8 mus musculus
11	696.5	35.6	367	11 Q8EHL1	Q8EHL1 mus musculus
12	694.5	35.5	369	4 Q9BYF7	Q9BYF7 homo sapien
13	687.5	35.1	387	11 Q8BG86	Q8BG86 mus musculus
14	686.5	35.0	387	11 Q9D1Q5	Q9D1Q5 mus musculus
15	684.5	34.9	377	11 Q88804	Q88804 mus musculus
16	684.5	34.9	385	11 Q8VHQ0	Q8VHQ0 mus musculus

17	682	34.8	388	11 Q9Z2G2	Q9Z2G2 mus musculus
18	663	33.8	386	13 Q73860	Q73860 meleagris g
19	660.5	33.7	389	11 Q8DCD0	Q8DCD0 mus musculus
20	658	33.6	386	11 Q8D1E7	Q8D1E7 mus musculus
21	655	33.4	338	4 Q9AYF8	Q9AYF8 homo sapien
22	645	32.9	374	11 Q9E797	Q9E797 mus musculus
23	642.5	32.8	375	11 Q8K3Y1	Q8K3Y1 mus musculus
24	634.5	31.9	377	11 Q9DAV6	Q9DAV6 mus musculus
25	619	31.6	374	11 Q88800	Q88800 mus musculus
26	613.5	31.3	377	11 Q8BMT0	Q8BMT0 mus musculus
27	609.5	31.1	397	11 Q8K1K6	Q8K1K6 mus musculus
28	606.5	31.0	377	11 Q88806	Q88806 mus musculus
29	601.5	30.7	379	11 Q9D0S8	Q9D0S8 mus musculus
30	593	30.3	450	13 Q8AYE3	Q8AYE3 brachydanio
31	589.5	30.1	377	11 Q8VHQ1	Q8VHQ1 mus musculus
32	588.5	30.0	397	11 Q8K3K4	Q8K3K4 rattus norv
33	585.5	29.9	453	13 Q8UVS2	Q8UVS2 struthio ca
34	583.5	29.8	377	11 Q9DAZ7	Q9DAZ7 mus musculus
35	579.5	29.6	456	13 Q8UVS0	Q8UVS0 xenopus lae
36	573.5	29.3	453	13 Q8UVS1	Q8UVS1 cheilydra se
37	570.5	29.1	452	13 Q9PTA8	Q9PTA8 salmo salar
38	570	29.1	426	17 Q8TNN7	Q8TNN7 methanosarc
39	568	29.0	410	13 Q73790	Q73790 gallus gall
40	565.5	28.9	423	13 Q91422	Q91422 gallus gall
41	564	28.8	465	4 Q13815	Q13815 homo sapien
42	555	28.3	359	11 Q9D6A7	Q9D6A7 mus musculus
43	553	28.2	459	11 Q9WTT1	Q9WTT1 cavia porce
44	550.5	28.1	424	17 Q8TKL5	Q8TKL5 methanosarc
45	542.5	27.7	448	13 Q9W648	Q9W648 fugu rubrip

ALIGNMENTS

RESULT 1

Q9D695	PRELIMINARY;	PRT;	380 AA.
ID Q9D695	AC Q9D695;		
DT 01-JUN-2001	(TREMELrel. 17, Created)		
DT 01-JUN-2001	(TREMELrel. 17, Last sequence update)		
DT 01-MAR-2003	(TREMELrel. 23, Last annotation update)		
DE 4631416M05Rik	protein (Megin).		
GN SERFINB7 OR 4631416M05Rik.			
OS Mus musculus (Mouse).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX NCBI_TaxID=10090;			
RN [1]_taxid=10090;			
RP SEQUENCE FROM N.A.			
RC STRAIN=C57BL/6J; TISSUE=Skin;			
RX MEDLINE=21085660; PubMed=11217851;			
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,			
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,			
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,			
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,			
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,			
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,			
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,			
RA Schriml L.M., Staubli P., Suzuki R., Tomita M., Wagner L., Washio T.,			
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,			
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaudo M.F.,			
RA Brownstein X.J., Buit C., Fletcher C., Fujita M., Gariboldi M.,			
RA Gustincich S., Hall D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,			
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,			
RA Nordore P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,			
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,			
RA Suzuki H., Tcyo-oka K., Wang X.H., Weitz C., Whittaker C., Wilming L.,			
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kotsuki S.,			
RA Hayashizaki Y.,			
RL "Functional annotation of a full-length mouse cDNA collection."			
RN Nature 409:685-690(2001).			
RP SEQUENCE FROM N.A.			

```

RX MEDLINE=23368006; PubMed=11473647;
RA Nangaku M., Miyata T., Suzuki D., Umezono T., Hashimoto T., Wada T.,
RA taga M., Nagano N., Inagi R., Kurokawa K.; regulation in
RT "Cloning of rodent meginin revealed its up-regulation in
RT mesangio proliferative nephritis.";
RL Kidney Int. 60:641-652(2001).
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EXBL: AK014524; BAB29410.1; -
DR HSSP: AF105328; AAL16768.1; -
DR HSSP: P05121; IA7C.
DR MGD: MGI:2151053; Serpinb7.
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; serpin; 1.
DR SMART: SM00093; SERPIN; 1.
DR PROSITE: PS00284; SERPIN; 1.
DR Protease inhibitor; Serine protease inhibitor; Serpin.
KW Protease inhibitor; Serine protease inhibitor; Serpin.
SQ SEQUENCE 380 AA; 43050 MW; C9240272BCFB9CF4 CRC64;

Query Match 76.8%; Score 1505; DB 11; Length 380;
Best Local Similarity 73.9%; Pred. No. 2.2e-99;
Matches 281; Conservative 53; Mismatches 46; Indels 0; Gaps 0;

QY 1 MASLAANAARFCNLFREMDNQGNVFPSSLSLPAALALVRLGAGQDSSLSQIDKLIHV 60
DB 1 MASLAANAARFCNLFREMDNQGNVFPSSLSLPAALALVRLGAGQDSSLSQIDKLIHV 60
QY 61 NTASGYGNSNSQSGQLQKVFSDINASHKDYDLSIVNGLPFAKYVGFHKDYIECAEK 120
DB 61 NTASGYGNSNSQSGQLQKVFSDINASHKDYDLSIVNGLPFAKYVGFHKDYIECAEK 120
QY 121 LYDAKVERVDTNHELDTRNINKWENETHGKIKNVIGEGGSSSAVMVLNAYVFKG 180
DB 121 LYDAKVERVDTNHELDTRNINKWENETHGKIKNVIGEGGSSSAVMVLNAYVFKG 180
QY 181 WSAFTKSETINCHFKSPCKSGKAVMMHQRKFNLSVIEDPSMKILELYNGGIMTVL 240
DB 181 WSAFTKSETINCHFKSPCKSGKAVMMHQRKFNLSVIEDPSMKILELYNGGIMTVL 240
QY 241 LPENDLSIENKLTFFQNLMEWTNPRMTSKYVEVFPFPQFKIEKNYEMKYLALGKIDIF 300
DB 241 LPENDLSIENKLTFFQNLMEWTNPRMTSKYVEVFPFPQFKIEKNYEMKYLALGKIDIF 300
QY 301 DESKADLSIAGSGRLYIRSMVHKSYIEVTEGTEATATGNSNIVEKQLPOSTLFRADHP 360
DB 301 DESKADLSIAGSGRLYIRSMVHKSYIEVTEGTEATATGNSNIVEKQLPOSTLFRADHP 360
QY 361 FLFVIRKDDIILFSGKVSCP 380
DB 361 FLFVIRKDDIILFSGKVSCP 380

RESULT 2
Q920J5 PRELIMINARY; PRT; 380 AA.
ID Q920J5
AC Q920J5
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE RIKEN cDNA 2310046M08 Gene.
GN 2310046M08RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EXBL: BC010313; AAL10313.1; -
DR HSSP: P01009; IQLP.
DR MGD: MGI:1914207; 2310046M08RIK.
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; serpin; 1.
DR SMART: SM00093; SERPIN; 1.
DR Protease inhibitor; Serine protease inhibitor; Serpin.
KW Protease inhibitor; Serine protease inhibitor; Serpin.
SQ SEQUENCE 380 AA; 43544 MW; AF4BB62FB084E6FA CRC64;

Query Match 40.8%; Score 800; DB 11; Length 388;
Best Local Similarity 41.8%; Pred. No. 4.5e-49;
Matches 162; Conservative 81; Mismatches 137; Indels 8; Gaps 5;

QY 1 MASLAANAARFCNLFREMDNQGNVFPSSLSLPAALALVRLGAGQDSSLSQIDKLIHV 60

```

```

DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; serpin; 1.
DR SMART: SM00093; SERPIN; 1.
DR PROSITE: PS00284; SERPIN; 1.
DR Protease inhibitor; Serine protease inhibitor; Serpin.
KW Protease inhibitor; Serine protease inhibitor; Serpin.
SQ SEQUENCE 380 AA; 42821 MW; DB076CA8EB2C2FBC CRC64;

Query Match 76.0%; Score 1488; DB 11; Length 380;
Best Local Similarity 73.7%; Pred. No. 3.6e-98;
Matches 280; Conservative 57; Mismatches 43; Indels 0; Gaps 0;

QY 1 MASLAANAARFCNLFREMDNQGNVFPSSLSLPAALALVRLGAGQDSSLSQIDKLIHV 60
DB 1 MASLAANAARFCNLFREMDNQGNVFPSSLSLPAALALVRLGAGQDSSLSQIDKLIHV 60
QY 61 NTASGYGNSNSQSGQLQKVFSDINASHKDYDLSIVNGLPFAKYVGFHKDYIECAEK 120
DB 61 NTASGYGNSNSQSGQLQKVFSDINASHKDYDLSIVNGLPFAKYVGFHKDYIECAEK 120
QY 121 LYDAKVERVDTNHELDTRNINKWENETHGKIKNVIGEGGSSSAVMVLNAYVFKG 180
DB 121 LYDAKVERVDTNHELDTRNINKWENETHGKIKNVIGEGGSSSAVMVLNAYVFKG 180
QY 181 WSAFTKSETINCHFKSPCKSGKAVMMHQRKFNLSVIEDPSMKILELYNGGIMTVL 240
DB 181 WSAFTKSETINCHFKSPCKSGKAVMMHQRKFNLSVIEDPSMKILELYNGGIMTVL 240
QY 241 LPENDLSIENKLTFFQNLMEWTNPRMTSKYVEVFPFPQFKIEKNYEMKYLALGKIDIF 300
DB 241 LPENDLSIENKLTFFQNLMEWTNPRMTSKYVEVFPFPQFKIEKNYEMKYLALGKIDIF 300
QY 301 DESKADLSIAGSGRLYIRSMVHKSYIEVTEGTEATATGNSNIVEKQLPOSTLFRADHP 360
DB 301 DESKADLSIAGSGRLYIRSMVHKSYIEVTEGTEATATGNSNIVEKQLPOSTLFRADHP 360
QY 361 FLFVIRKDDIILFSGKVSCP 380
DB 361 FLFVIRKDDIILFSGKVSCP 380

RESULT 3
Q921L2 PRELIMINARY; PRT; 388 AA.
ID Q921L2
AC Q921L2
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE RIKEN cDNA 2310046M08 Gene.
GN 2310046M08RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EXBL: BC010313; AAL10313.1; -
DR HSSP: P01009; IQLP.
DR MGD: MGI:1914207; 2310046M08RIK.
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; serpin; 1.
DR SMART: SM00093; SERPIN; 1.
DR Protease inhibitor; Serine protease inhibitor; Serpin.
KW Protease inhibitor; Serine protease inhibitor; Serpin.
SQ SEQUENCE 388 AA; 43544 MW; AF4BB62FB084E6FA CRC64;

Query Match 40.8%; Score 800; DB 11; Length 388;
Best Local Similarity 41.8%; Pred. No. 4.5e-49;
Matches 162; Conservative 81; Mismatches 137; Indels 8; Gaps 5;

QY 1 MASLAANAARFCNLFREMDNQGNVFPSSLSLPAALALVRLGAGQDSSLSQIDKLIHV 60

```



```

Db 111 MYGAOLAPVDFPHASEDARKEINQWVKQTEGKIPELLSVGVWDSMTKLVLNVAIYFKGM 170
QY 181 WQSAATKSETINCHFKSPKCSKAVAMVHQRKFNLSVIEDPSWKILEURYGG-INAYV 239
Db 171 WEEKFWTEDITDAPFLSKKOTRTVKWYQKXKPGYISDLCKCKVLEMPYQGGELSMVI 230
QY 240 LLPEN-----DASEIENKLTQNLMEWTPRMTSKYVEVFFPQFKIEKNYEMQYLRA 293
Db 231 LPPKIDIESNGLKLEKQILEKLEWTRNLEFIDVHVKLPFKFIESYLSNIGR 290
QY 294 IGLKIDFDSKADLSGIASGRGLYISRMHKSIVETEGTEATAATGNSIVEXQLPOST 353
Db 291 IGVQDLSSKADLSGWSGRDLPTFSKIVHKSFEVNEEGTEAAATGSIATFCMLPDEE 350
QY 354 LFRADHPFLFVKIDDI--ILFSKGVSCP 380
Db 351 ERTVDHPFIFIRHNPTSNVLFLGRVSCP 379

RESULT 11
QB8HL1
ID QB8HL1 PRELIMINARY; PRT; 387 AA.
AC QB8HL1
DT 01-MAR-2003 (TRENHRel. 23, Created)
DT 01-MAR-2003 (TRENHRel. 23, Last sequence update)
DE Squamous cell carcinoma antigen 2 related protein 1.
GN SCA2-RS1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
XP SEQUENCE FROM N.A.
RA STRAIN=Car-S, and Car-R; TISSUE=TPA-treated skin;
RA Gariboldi M., Peissel B., Fabbri A., Saran A., Zaffaroni D.,
RA Spicola M., Falvella P.S., Pazzaglia S., Tanuma J.-I., Maurichi A.,
RA Bartoli C., Silverman G., Coveilli V., Piloti S., Hayashiaki Y.,
RA Okazaki Y., Dragani T.A.;
RT "The serpin Sca2 gene plays a functional role in genetic
RT susceptibility to skin tumorigenesis in mice and humans."
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY144685; AAN62872.1; -.
DR EMBL; AY144686; AAN62873.1; -.
SQ SEQUENCE 367 AA; 44539 MW; 0AD6398FA0251135 CRC64;

Query Match 35.6%; Score 696.5; DB 11; Length 387;
Best Local Similarity 37.3%; Pred. No. 1.1e-41;
Matches 145; Conservative 91; Mismatches 142; Indels 11; Gaps 6;

QY 1 MASLAANAECFNLFRWDDNQGNGVFPSSLSLFAALALVRLGQDDSLQIDKLH- 59
Db 1 MIRHAADKFAVEMVQL--RESQKNIFYSPISTMTALAWLQLGAKNGTQIERKVLQF 58
QY 60 VNTASGVGNS--NSQSGLOSQKRVFSDINASHKDYDLSVNGLFAKVGYGHKDYIE 116
Db 59 IETTKTKTKSEHCDDEENVHEQFKLITQLKNSDDYDLKAANSYGAQFPFLQFLE 118
QY 117 CAELKYDAKVERVDFTNHLBETRNINKWVENETHGKIKNVIGEGGSSAVMLVNAVY 176
Db 119 DIKRYQAKVESLDFEATERSSEKINSVESKTKGKIDLPFGSGSLSSSTILVNAVY 178
QY 177 FKGWQSAFTKSETINCHFKSPKCSKAVAMVHQRKFNLSVIEDPSWKILEURYNG-GI 235
Db 179 FKGWNAKDEHNTREKFWLNKNTSKPYQMMKQKRNKFNFSFLGVHQAIVSEIYKQKOL 238
QY 236 NMVYLLPE--NDLSIENKLTQNLMEWTPRMTSKYVEVFFPQFKIEKNYEMQYLRA 293
Db 239 SMFVLLPMEIDGKLEQLQTLKLEWTRNLEFIDVHVKLPFKFIESYLSNIGR 298
QY 294 IGLKIDFDSKADLSGIASGRGLYISRMHKSIVETEGTEATAATGNSIVEXQLPOST 353
Db 299 MGVMDAPDQKADPFGSMSSIPGLWSKVLHKSFEVNEEGTEAAATGVEVSQAIAE 358

```

```

QY 354 LFRADHPFLFVI--RKDDIILFSKGVSCP 380
Db 359 DFCDHPFLFIIHRKTNLSLFGRISSP 387

RESULT 12
Q9BYF7
ID Q9BYF7 PRELIMINARY; PRT; 369 AA.
AC Q9BYF7
DT 01-JUN-2001 (TRENHRel. 17, Created)
DT 01-JUN-2001 (TRENHRel. 17, Last sequence update)
DT 01-OCT-2002 (TRENHRel. 22, Last annotation update)
DE SCA2B.
GN SCA2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Suminami Y., Kishi F., Murakami A., Sakaguchi Y., Kato H.;
RT "Novel Forms of SCC Antigen Transcripts Produced by Alternative
RT Splicing."
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL; AB046400; BAB40773.1; -.
DR HSSP; P01008; 1ATH.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; Serpin; 1.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PSC0284; SERPIN; 1.
DR Protease inhibitor; Serine protease inhibitor; Serpin.
SQ SEQUENCE 369 AA; 42287 MW; CED7A2426378DE3E CRC64;

Query Match 35.5%; Score 694.5; DB 4; Length 369;
Best Local Similarity 39.0%; Pred. No. 1.4e-41;
Matches 153; Conservative 82; Mismatches 122; Indels 35; Gaps 10;

QY 1 MASLAANAECFNLFRWDDNQGNGVFPSSLSLFAALALVRLGQDDSLQIDKLH- 60
Db 1 MNSLSEANTKPMFLDLPQFRKSKEN-NIFYSPISITSLGMLVGLGAKNTAQQISKVLHF 59
QY 61 -----NTASGVGNSQSG-LQSLXRVFSDINASHKDYDLSVNGLFAKVGYGHKDY 114
Db 60 DQWTEHTERKAATYHVDRSGNVHGFQKLLATEBNKSTDAYELKIANKLFGKTYQLQFY 119
QY 115 IECAELKYDAKVERVDFTNHLBETRNINKWVENETHGKIKNVIGEGGSSAVMLVNA 174
Db 120 LDAIKKPYQTSVSTDFANAPEESRKKINSWESQTNKIKNLPDQGTGNDTTLVLNA 179
QY 175 VYFKGWQSAFTKSETINCHFKSPKCSKAVAMVHQRKFNLSVIEDPSWKILEURYNG- 233
Db 180 IYFKQWENFKENT-----KEEFWPN--KDVQAKVLEIPIYKKG 218
QY 234 GINMVTLLPE--NDLSIENKLTQNLMEWTPRMTSKYVEVFFPQFKIEKNYEMQYL 291
Db 219 DLSMVTLLPNEIDGLQKLEELTAELKLEWTSIQNRETCDLHLPRFWEESYDLKYL 278
QY 292 RALGKIDFDSKADLSGIASGRGLYISRMHKSIVETEGTEATAATGNSIVEXQLPQ 351
Db 279 RTMGWNAIFN-GPADLSGMTWSHGLSVSKVLHKAFEVTEEGVEAAAAATAVVVVELSSPS 337
QY 352 SP-LFRADHPFLFVIR--KDDIILFSKGVSCP 380
Db 338 TNEEFCCNHEFFLFFIRQNKTNLSILFYGRFSP 369

RESULT 13
QB8G86
ID QB8G86 PRELIMINARY; PRT; 387 AA.
AC QB8G86
DT 01-MAR-2003 (TRENHRel. 23, Created)

```

```

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Squamous cell carcinoma antigen 2.
GN SERPINB4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Car-S, and Car-B; TISSUE=TPA-treated skin;
RA Gariboldi M., Peissel B., Fabbri A., Saran A., Zaffaroni D.,
RA Spinoia M., Falvello F.S., Fazzaglia S., Tanuma J.-I., Maurichi A.,
RA Bartoli C., Silverman G., Covelli V., Pilotti S., Hayashizaki Y.,
RA Okazaki Y., Dragani T.A.;
RT "The serpin Sccl2 gene plays a functional role in genetic
RT susceptibility to skin tumorigenesis in mice and humans.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY144683; AAK2870.1; -.
DR EMBL; AY144684; AAK2871.1; -.
SQ SEQUENCE 387 AA; 44434 MW; B5B15F807FEF0688 CRC64;

Query Match 35.1%; Score 687.5; DB 11; Length 387;
Best Local Similarity 37.0%; Pred. No. 4.8e-41;
Matches 145; Conservative 87; Mismatches 143; Indels 17; Gaps 7;

QY 1 MASLAANAACFCNLFREMDNQGNGNVPFSSLSLFAALALVRLGACQDLSQIDKLLHV 60
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 1 MHLFABATTKFTLELYRQL--RESQNFYSPISMTALAMQLGAKGNTKQIEKVLQF 58
QY 61 NFAAGYGNSSNS---OSGLQSLKRVFSDINASHKDYLSIVNGLFAEKVYGFHKDYE 116
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 59 NETTKTKTEKSHCHDENVHEQFQKLTQLNKSNDYDLKAANSIYGAKGFPFVQTFLE 118
QY 117 CAEKLYDAKVERVDFTHLEEDTRRNKWNETHGKIKNVIGEGC:SSSAVMVLNAVY 176
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 119 DIKEYYQANVESLDFEAAESEKINSWESQNGKIDLPNGSLNRSTIILVNAVY 178
QY 177 FKQWOSAFKSTETINCHFKSPKCGKAVAMMHQERKFNLSVIEDPSMKILELYNG-GI 235
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 179 FKQWNRKFNENHTEEEKFWLNKNTSKPQWKQKRNKFNPSFLGDVHAQIVEIPYKQDL 238
QY 236 NMVLLPE--NDLSEINKLTFQNLMEWNTNPRMTSKYVEVFPQFKIKNYEMKQYLRA 293
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 239 SMVLVPEINGLKQLEBEQTLDKLEWTRANNMHMTLYLSLPFPKVDKDYDLPFLEH 298
QY 294 LGLKIDFDESKADLSGIASGRILYISRMWHKSYIEVTEGTEATATGGSNIVEKQ:POST 353
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 299 MGWVDAFDPQKADFGSMSTQGLVSKVLHKSFEVYNEGTEAATG--VEVSLTSAQ 355
QY 354 L---FRADHPFLFVI--RKDDILFSGKVSCP 380
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 356 IAEDFCCDHPFLFIIRKTNLSILFGRISPP 387

RESULT 14
Q9D1Q5 ID Q9D1Q5 PRELIMINARY; PRT; 387 AA.
AC Q9D1Q5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE 18 days embryo cDNA, Riken full-length enriched library,
DE clone:1110001H02, full insert sequence.
GN SCCA2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RC MEDLINE=21085660; PubMed=11217851;

```

```

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staib F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.P.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guskinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momotets P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL; AK003220; BAB22650.1; -.
DR HSSP; P05120; 1BY7.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; serpin; 1.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
KW Protease inhibitor; Serine protease inhibitor; Serpin.
SQ SEQUENCE 387 AA; 44558 MW; AE25C02520F3B194 CRC64;

Query Match 35.0%; Score 686.5; DB 11; Length 387;
Best Local Similarity 36.8%; Pred. No. 5.6e-41;
Matches 143; Conservative 91; Mismatches 144; Indels 11; Gaps 6;

QY 1 MASLAANAACFCNLFREMDNQGNGNVPFSSLSLFAALALVRLGACQDLSQIDKLLH- 59
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 1 MIRPHADVKFAVENYRQL--RESQNFYSPISMTALAMQLGAKGNTKQIEKVLQF 58
QY 60 VNTASGYGNSS---NSQSGIQLKRVFSDINASHKDYLSIVNGLFAEKVYGFHKDYE 116
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 59 IETTKTKTEKSHCHDENVHEQFQKLTQLNKSNDYDLKAANSIYGAKGFPFVQTFLE 118
QY 117 CAEKLYDAKVERVDFTHLEEDTRRNKWNETHGKIKNVIGEGC:SSSAVMVLNAVY 176
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 119 DIKEYYQANVESLDFEAAESEKINSWESQNGKIDLPNGSLNRSTIILVNAVY 178
QY 177 FKQWOSAFKSTETINCHFKSPKCGKAVAMMHQERKFNLSVIEDPSMKILELYNG-GI 235
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 179 FKQWNRKFNENHTEEEKFWLNKNTSKPQWKQKRNKFNPSFLGDVHAQIVEIPYKQDL 238
QY 236 NMVLLPE--NDLSEINKLTFQNLMEWNTNPRMTSKYVEVFPQFKIKNYEMKQYLRA 293
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 239 SMFVLLPEIDGLKQLEBEQTLDKLEWTRANNMHMTLYLSLPFPKVDKDYDLPVLSH 298
QY 294 LGLKIDFDESKADLSGIASGRILYISRMWHKSYIEVTEGTEATATGGSNIVEKQ:POST 353
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 299 MGWVDAFDPQKADFGSMSTQGLVSKVLHKSFEVYNEGTEAATGVEVSVRSQAIAE 358
QY 354 LFRADHPFLFVI--RKDDILFSGKVSCP 380
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 359 DFCCDHPFLFIIRKTNLSILFGRICSP 387

RESULT 15
Q08804 ID Q08804 PRELIMINARY; PRT; 377 AA.
AC Q08804;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Serine protease inhibitor NK13 (Serine protease inhibitor 12).
GN SP112.
OS Mus musculus (Mouse).

```


